The Critical Attractive Random Polymer in Dimension One

Remco van der Hofstad[1], Achim Klenke[2], Wolfgang Konig
Basic Research Institute in the Mathematical Sciences
HP Laboratories Bristol
HPL-BRIMS-2001-05
March 16th , 2001*

email: R.W. vanderHofstad@its.tudelft.nl, klenke@mi.uni-erlangen.de, koenig@math.tu-berlin.de

one-dimensional polymers, repulsive and attractive interaction, phase transition, local times of simple random walk

A polymer chain with attractive and repulsive forces between the monomers is modeled by attaching a weight e^{-b} for every self-intersection and $e^{\gamma/(2d)}$ for every self-contact to the probability of an n-stepsimple random walk on $\mathbb{Z}^d$, where $b, \gamma > 0$ are parameters. It is known that for $d = 1$ and $\gamma > b$ the chain collapses down to finitely many sites, while for $d = 1$ and $\gamma < b$ it spreads out ballistically. Here we study for $d = 1$ the critical case $\gamma = b$ and show that the end-to-end distance runs on the scale $\alpha_n = \sqrt{n}(\log n)^{-1/4}$. We describe the asymptotic shape of the accordingly scaled local times in terms of an explicit variational formula and prove that the scaled polymer chain occupies a region of size α_n times a constant. Moreover, we derive the asymptotics of the partition function.

* Internal Accession Date Only Approved for External Publication
[1] Department of Applied Mathematics, Delft University of Technology, 2628 CD Delft, The Netherlands
[2] Department of Mathematics, Erlangen-Nurnberg University, Bismarckstraβe 1 ½, 91054 Erlangen, Germany

The critical attractive random polymer in dimension one

Remco van der Hofstad* Achim Klenke† Wolfgang König‡,§

February 20, 2001

Abstract

A polymer chain with attractive and repulsive forces between the monomers is modeled by attaching a weight $e^{-\beta}$ for every self-intersection and $e^{\gamma/(2d)}$ for every self-contact to the probability of an n-step simple random walk on $\mathbb{Z}^d$, where $\beta, \gamma > 0$ are parameters. It is known that for $d = 1$ and $\gamma > \beta$ the chain collapses down to finitely many sites, while for $d = 1$ and $\gamma < \beta$ it spreads out ballistically.

Here we study for $d = 1$ the critical case $\gamma = \beta$ and show that the end-to-end distance runs on the scale $\alpha_n = \sqrt{n}(\log n)^{-1/4}$. We describe the asymptotic shape of the accordingly scaled local times in terms of an explicit variational formula and prove that the scaled polymer chain occupies a region of size α_n times a constant. Moreover, we derive the asymptotics of the partition function.

MSC 2000. 60F05, 60J15, 60J55.

Keywords and phrases. One-dimensional polymers, repulsive and attractive interaction, phase transition, Knight's theorem for local times of simple random walk.

0 Introduction and main results

Introduction

Polymers are large molecules that are built of smaller units. These smaller units are either all of the same type or of two or at most a small number of different types. Typically these building blocks allow two chemical bonds (of fixed length, e.g., 1.5×10^{-10}m for polyethylene) to neighboring monomers. Hence a polymer is typically a linear structure. The stereometric angels between the bonds may vary. Thus the geometry of a polymer is typically quite complicated. Under thermic influences it is even random.

A fundamental quantity in both the experimental and theoretical study of polymers is the quantitative connection between the number of monomers in a polymer chain and the radius of gyration which is a measure for the spatial extent of the polymer. For the theoretical study of polymers the chemist

*Department of Applied Mathematics, Delft University of Technology, 2628 CD Delft, The Netherlands. R.W.vanderHofstad@its.tudelft.nl

†Department of Mathematics, Erlangen-Nürnberg University, Bismarckstraße $1\frac{1}{2}$, 91054 Erlangen, Germany. klenke@mi.uni-erlangen.de

‡BRIMS, Hewlett-Packard Laboratories, Filton Road, Stoke Gifford, Bristol BS12 6QZ, United Kingdom. koenig@math.tu-berlin.de

§on leave from TU Berlin, Germany

P.J. Flory suggested (see [Fl71]) the following stochastic model as a caricature of a polymer with $n+1$ monomers: A polymer corresponds to a random walk path $(S_0, \dots, S_n)$ in $\mathbb{Z}^d$ with law P. Attractive and repulsive forces between monomers (other than the forces of the direct bonds), are modeled by a Hamiltonian (energy function) H_n which is a function of the path $(S_0, \dots, S_n)$. The distribution Q_n of the random polymer is then the Boltzmann distribution defined by

$$\frac{dQ_n}{dP} = Z_n^{-1} e^{-H_n}, \tag{0.1}$$

where $Z_n = E(e^{-H_n})$ is the *partition function* (normalizing constant). (Expectations w.r.t. P are denoted by E.)

For an expository paper on mathematical polymer models, see [dH96]. For a survey paper on results for one-dimensional polymers, see [HKö01]. For an introduction to polymers from a physicist's point of view, see [Va98].

In this paper we consider the situation where the Hamiltonian $H_n^{\beta,\gamma}$ depends on two parameters $\beta, \gamma \in [0, \infty)$ and is defined by

$$H_n^{\beta,\gamma} = \beta \sum_{i,j=0}^{n} 1\!\!1\{S_i = S_j\} - \frac{\gamma}{2d} \sum_{i,j=0}^{n} 1\!\!1\{|S_i - S_j| = 1\}. \tag{0.2}$$

In words, $H_n^{\beta,\gamma}$ is equal to β times the number of self-intersections minus γ times the number of self-contacts by time n.

The law $Q_n^{\beta,\gamma}$ is called the *n-polymer measure with strength of repellence β and strength of attraction γ*. It gives a penalty $e^{-2\beta}$ to every pair of monomers at the same site and a reward $e^{\gamma/d}$ to every pair of neighboring monomers. The penalty models polarization of the monomers, the so-called excluded-volume effect. The reward models the situation in which there are attractive forces between the monomers.

The main goal of this article is to study this model for simple random walk in dimension $d = 1$ for $\beta = \gamma$ where a phase transition in the asymptotic behavior of the spatial extension occurs.

Conjectures and earlier results

It is generally believed that, under $Q_n^{\beta,\gamma}$, the path exhibits asymptotic behavior drastically different from the diffusive behavior of simple random walk. More precisely, one conjectures that

$$E_{Q_n^{\beta,\gamma}}(|S_n|) \approx n^{\nu(\beta,\gamma)}, \qquad n \to \infty, \tag{0.3}$$

with some characteristic exponent $\nu(\beta, \gamma) \in [0, 1]$. There is no problem in defining the measure $Q_n^{\infty,\gamma}$ analogously, hence we are going to assume that $\beta \in (0, \infty]$.

In dimension one, it is shown in [HKl01] that for $\gamma < \beta$, the polymer behaves ballistically, in the sense that the number of sites grows like n. This identifies $\nu(\beta, \gamma) = 1$ for $\gamma < \beta$. Moreover, it is shown that for $\gamma > \beta$, the polymer collapses to a finite number of points, so that $\nu(\beta, \gamma) = 0$ for $\gamma > \beta$. In this paper, we will investigate the critical case $\gamma = \beta$. In dimensions greater than one, there is a richer structure. Indeed, there it is expected to have two phase transition. The asymptotic behavior of the polymer is expected to have three possibilities: a collapse to a finite number of points for $\gamma > \beta$, a dense packing of building blocks so that $\nu = 1/d$ for intermediate γ, and self-avoiding walk behavior for $\gamma \ll \beta$. For a more detailed description of the conjectures in dimension $d > 1$, see [HKl01].

Intuitively, we can explain these conjectures and results as follows. It is helpful to rewrite the Hamiltonian in terms of the walker's so-called local times. Define

$$\ell_n(x) = \sum_{i=0}^{n} 1\!\!1\{S_i = x\}, \qquad n \in \mathbb{N}_0, x \in \mathbb{Z}^d, \tag{0.4}$$

the number of monomers at x of the n-polymer chain. Then we have the identity

$$H_n^{\beta,\gamma} = (\beta - \gamma) \sum_{x \in \mathbb{Z}^d} \ell_n(x)^2 + \frac{\gamma}{4d} \sum_{x \in \mathbb{Z}^d, e \sim 0} [\ell_n(x) - \ell_n(x+e)]^2. \tag{0.5}$$

where the sum over $e \sim 0$ runs over all the $2d$ neighbors e of the origin.

One has to examine the path's best strategy to minimize $H_n^{\beta,\gamma}$ in (0.5) without having too little probability under P. If $\gamma > \beta$, then it is most favorable for the polymer to stay in a bounded region, so that the Hamiltonian is of order $-n^2$. If $\gamma \ll \beta$, then it is most favorable to avoid self-intersections, and hence the behavior is similar to the behavior of self-avoiding walk. If $0 \ll \gamma < \beta$ and $\beta - \gamma$ is small, then the penalty for large nearest-neighbor local time differences is much larger than the penalty for self-intersections. Hence, it is most favorable for the polymer to minimize these differences. The only way to do that, and keep the sum of squares of local times of the order n is for the polymer to clump together in a region of the order $n^{1/d}$.

Description of the results

In this paper we will study the critical case where $\beta = \gamma$ in dimension one. For the remainder of the paper, fix $d = 1$ and assume that $(S_n)_{n \in \mathbb{N}_0}$ is an ordinary simple random walk on $\mathbb{Z}$ starting at $S_0 = 0$.

Observe from (0.5) that our Hamiltonian is given in terms of the local times as

$$H_n^{\beta,\beta} = \frac{\beta}{2} \sum_{x \in \mathbb{Z}} [\ell_n(x) - \ell_n(x+1)]^2 . \tag{0.6}$$

Thus, the path measure $Q_n^{\beta,\beta}$ is concentrated on paths whose local times are close together in neighboring sites, but there is no explicit repulsion or attraction effect.

In order to describe our result, we have to introduce some notation. Along the way, we give an informal description of our result. The precise statement appears in Theorem 1 below.

It turns out that, under $Q_n^{\beta,\beta}$, the endpoint S_n of the polymer chain runs on scale

$$\alpha_n = \frac{n^{1/2}}{(\log n)^{1/4}}, \qquad n \in \mathbb{N} \setminus \{1\}. \tag{0.7}$$

The accordingly scaled continuous version $\overline{\ell}_n \colon \mathbb{R} \to [0, \infty)$ of the local times is defined by

$$\overline{\ell}_n\left(\frac{x}{\alpha_n}\right) = \frac{\alpha_n}{n} \ell_n(x), \qquad x \in \mathbb{Z}, \tag{0.8}$$

and by linear interpolation between the points in $\mathbb{Z}/\alpha_n$. Note that $\overline{\ell}_n$ is an element of the set

$$\mathcal{F} = \left\{ \varphi \in \mathcal{AC}_c(\mathbb{R}, [0, \infty)) \colon \int_{\mathbb{R}} \varphi(t)\, dt = 1 \right\} \tag{0.9}$$

of absolutely continuous, compactly supported Lebesgue densities on $\mathbb{R}$. The asymptotics of the partition function $Z_n^{\beta,\beta}$ turns out to be determined by the functional $\mathcal{G}_\beta \colon \mathcal{F} \to [0, \infty)$ given by

$$\mathcal{G}_\beta(\varphi) = \frac{\beta}{2} \int_{\mathbb{R}} \varphi'(t)^2\, dt + \frac{1}{4} |\operatorname{supp}(\varphi)|. \tag{0.10}$$

A fundamental role in this paper is played by the variational problem connected with $\mathcal{G}_\beta$. We define

$$\chi_\beta = \inf \{ \mathcal{G}_\beta(\varphi) \colon \varphi \in \mathcal{F} \} . \tag{0.11}$$

We define the function $\varphi^\star_\beta \in \mathcal{F}$ by

$$\varphi^\star_\beta(t) = \tfrac{3}{4R_\beta\star}\left(1 - \left(\tfrac{t}{R^\star_\beta}\right)^2\right)_+, \qquad \text{where} \qquad R^\star_\beta = (\tfrac{9}{2}\beta)^{1/4}. \tag{0.12}$$

Our first partial result is that $\varphi^\star_\beta$ uniquely (up to shifts) minimizes $\mathcal{G}_\beta$.

Proposition 0.1 *The minimizer of $\mathcal{G}_\beta$ on $\mathcal{F}$ is unique up to translations and is equal to $\varphi^\star_\beta$, and the value of the minimum is*

$$\chi_\beta = \frac{\sqrt[4]{8}}{\sqrt{3}}\,\beta^{1/4}. \tag{0.13}$$

Our main result consists of three statements about the asymptotics of the random polymer. The first one is the identification of the logarithmic asymptotics of the partition sum $Z_n^{\beta,\beta}$ in terms of χ_β. The second statement is that $\overline{\ell}_n$ approaches a possibly random shift $\tau_\xi\varphi^\star_\beta$ of $\varphi^\star_\beta$ in the sense of the norm $\|\cdot\| = \|\cdot\|_1 + \|\cdot\|_\infty$. (We define $\tau_\xi\varphi(t) = \varphi(t-\xi)$ for $t,\xi \in \mathbb{R}$.) The third statement is that the Lebesgue measure of the support of $\overline{\ell}_n$, denoted by $|\operatorname{supp}\overline{\ell}_n|$, converges to the one of $\varphi^\star_\beta$, which is $2R^\star_\beta$. This implies in particular that the offset ξ is concentrated on $[-R^\star_\beta, R^\star_\beta]$. Here is the precise statement.

Theorem 1 *Fix $\beta > 0$ and put $d = 1$. Then*

(i)

$$\lim_{n\to\infty} \frac{1}{\alpha_n \log n} \log Z_n^{\beta,\beta} = -\chi_\beta. \tag{0.14}$$

(ii) For any $\varepsilon > 0$,

$$\lim_{n\to\infty} Q_n^{\beta,\beta}\Big(\inf_{\xi\in\mathbb{R}} \|\overline{\ell}_n - \tau_\xi\varphi^\star_\beta\| > \varepsilon\Big) = 0. \tag{0.15}$$

(iii) For any $\delta > 0$,

$$\lim_{n\to\infty} Q_n^{\beta,\beta}\big(\big||\operatorname{supp}\overline{\ell}_n| - 2R^\star_\beta\big| > \delta\big) = 0. \tag{0.16}$$

Quantitative statements on the rate of convergence can be found in Proposition 1.2 and 1.3 below.

Heuristics for the variational problem

Let us roughly calculate the contribution to $e^{-H_n^{\beta,\beta}}$ coming from paths satisfying $\overline{\ell}_n \approx \varphi$ for some fixed $\varphi \in \mathcal{F}$. If

$$\frac{1}{n}\ell_n(x) \approx \frac{1}{\alpha_n}\varphi\left(\frac{x}{\alpha_n}\right), \qquad x \in \mathbb{Z}, \tag{0.17}$$

for some sequence α_n in $(0,\infty)$ and some $\varphi \in \mathcal{F}$, then our Hamiltonian is roughly given by

$$H_n^{\beta,\beta} \approx \frac{\beta}{2}\frac{n^2}{\alpha_n^3}\int \varphi'(t)^2\,dt. \tag{0.18}$$

In order to obtain an approximation for the probability of the event in (0.17), we use the fact (Knight's Theorem) that $(\frac{1}{2}\ell_n(x))_{x\in\mathbb{Z}}$ is close to being a Markov chain on $\mathbb{N}$ with transition kernel

$$P(i,j) = \binom{i+j-2}{i-1}\left(\frac{1}{2}\right)^{i+j-1}, \qquad i,j \in \mathbb{N}. \tag{0.19}$$

This kernel has the limiting behavior

$$P(i,j) \approx \frac{1}{\sqrt{2\pi(i+j)}} \exp\left\{-\frac{(i-j)^2}{2(i+j)}\right\}, \qquad i,j \to \infty. \tag{0.20}$$

If we make the mild assumption that $\alpha_n = o(n)$, i.e., that the local times tend to infinity as $n \to \infty$, then we may use the asymptotics in (0.20). The term coming from the exponential in (0.20) is always negligible in comparison to the Hamiltonian in (0.18), and therefore we may concentrate on the square root term. Therefore, the leading term in the expansion is:

$$P\left(\overline{\ell}_n \approx \varphi\right) \approx \prod_{x\in\mathbb{Z}} P\left(\frac{n}{2\alpha_n}\varphi\left(\frac{x-1}{\alpha_n}\right), \frac{n}{2\alpha_n}\varphi\left(\frac{x}{\alpha_n}\right)\right) \approx \exp\left\{-\frac{1}{2}|\operatorname{supp}(\varphi)|\alpha_n \log\frac{n}{\alpha_n}\right\}. \tag{0.21}$$

Now one must choose the order of α_n in such a way that the order of the Hamiltonian in (0.18) and the order of the logarithm of the probability in (0.21) are equal, i.e.,

$$\frac{n^2}{\alpha_n^3} = \alpha_n \log\frac{n}{\alpha_n}. \tag{0.22}$$

This is the case precisely for the choice in (0.7). Since $\log(n/\alpha_n) \approx \frac{1}{2}\log n$, we end up with the formula

$$E\left(e^{-H_n^{\beta,\beta}} 1\!\!1\{\overline{\ell}_n \approx \varphi\}\right) \approx e^{-\mathcal{G}_\beta(\varphi)\alpha_n \log n}, \tag{0.23}$$

by combining the approximations (0.18) and (0.21). Maximization over $\varphi \in \mathcal{F}$ yields (0.14).

Discussion

In Theorem 1, the path measure $Q_n^{\beta,\beta}$ turns out to be asymptotically slightly self-attractive.

We have a precise conjecture about the limiting joint distribution of the random shift arising in assertion (ii) and the scaled endpoint $\frac{1}{\alpha_n}S_n$ of the polymer chain. In order to describe it, we need the square root of $\varphi_\beta^\star$ to be normalized as a probability density:

$$\mu_\beta^\star(d\xi) = \frac{1}{N_\beta^\star}\sqrt{\varphi_\beta^\star(\xi)}\,d\xi, \qquad \text{where} \quad N_\beta^\star = \int_{\mathbb{R}} \sqrt{\varphi_\beta^\star(\xi)}\,d\xi. \tag{0.24}$$

Let Y_β and Y'_β be independent random variables with law $\mu_\beta^\star$. Then we conjecture that

$$\mathcal{L}_{Q_n^{\beta,\beta}}[(S_n/\alpha_n, \overline{\ell}_n)] \overset{n\to\infty}{\Longrightarrow} \mathcal{L}[(Y_\beta + Y'_\beta, \tau_{Y_\beta}\varphi_\beta^\star)]. \tag{0.25}$$

(Here "$\mathcal{L}$" denotes the law of a random variable and "$\Longrightarrow$" denotes weak convergence.) This conjecture is based on heuristic calculations using the formulas in Section 3. Due to insufficient control on these formulas, which would go down to finite order asymptotics, we have not been able to prove (0.25).

The assertion (0.25) is typical for self-attractive models investigated in the literature. One of the most prominent examples is the discrete version of the Wiener-sausage model where one uses the Hamiltonian $H_n = \#\{S_0, \dots, S_n\}$ in an arbitrary dimension (see the monograph [Sz98] and the references therein). There also a variational problem arises whose solution is unique up to shifts, and the square root of the minimizer serves as an asymptotic density for the endpoint distribution. In self-attractive models, it is often possible to construct a transformed process in terms of which (0.25) can be proven. This technique, however, seems to fail in the present context.

The greatest difference between the Wiener sausage and the model in the present paper is that our variational formula does not rely on the Donsker-Varadhan large-deviation functional for the local times. Also, our functional is not genuinely self-attractive, and therefore the standard techniques of folding or periodization do not work here. Indeed, the upper bound in (0.14) is derived by a careful combinatorial analysis rather than by a compactification procedure.

It is easy to see from (0.25) that there should be a positive constant C such that, for every $\beta > 0$,

$$E_{Q_n^{\beta,\beta}}(|S_n|) \sim C\beta^{1/4}\alpha_n, \qquad n \to \infty. \tag{0.26}$$

(Our actual Theorem 1 only yields '$\leq$'.) This means that there would be a non-monotonicity in β at $\beta = 0$ for large n since $\alpha_n = o(\sqrt{n})$. We have no intuitive explanation for this to happen. Note that our "local times rate functional" $\varphi \mapsto \frac{1}{4}|\operatorname{supp}(\varphi)|$ favors highly concentrated φ's while the Donsker-Varadhan functional $\varphi \mapsto \int \varphi'(t)^2/\varphi(t)\,dt$ does not.

Organization of the paper

The rest of the paper is devoted to the proof of Proposition 0.1 and Theorem 1. In Section 1 we formulate three main assertions which imply our main results. The first one (Proposition 1.1) which is purely analytic is proved in Section 2, the other two (Proposition 1.2 and 1.3) are proved in Sections 4 and 5, respectively. Preparatory material for the latter sections is provided in Section 3.

Throughout the remainder of the paper, we fix $\beta \in (0,\infty)$ and suppress the dependence on β from the notation; in particular we shall write Q_n, Z_n, $\varphi^\star$ etc. instead of $Q_n^{\beta,\beta}$, $Z_n^{\beta,\beta}$, $\varphi_\beta^\star$ etc.

1 Strategy of the proof of Theorem 1

In this section we formulate three propositions from which Proposition 0.1 and Theorem 1 follow.

The first proposition is a stronger version of Proposition 0.1 and says in a strong sense that $\varphi^\star$ in (0.12) is the unique minimizer of $\mathcal{G}$ on $\mathcal{F}$. We denote the set of minimizers of $\mathcal{G}$ on $\mathcal{F}$ by

$$\mathcal{M} = \{\varphi \in \mathcal{F} \colon \mathcal{G}(\varphi) = \chi\}. \tag{1.1}$$

(Recall that we suppress the dependence on β from the notation.) Recall that $\tau_\xi\varphi$ denotes the translation of φ by $\xi \in \mathbb{R}$. The proposition formulates that $\mathcal{G}$ is bounded away from the minimal value, uniformly in the distance $\operatorname{dist}(\cdot,\mathcal{M})$ from the set $\mathcal{M}$. Here $\operatorname{dist}(\cdot,\mathcal{M})$ denotes the distance w.r.t. the norm $\|\cdot\| = \|\cdot\|_1 + \|\cdot\|_\infty$.

Proposition 1.1 (Variational problem)

(i)

$$\mathcal{M} = \{\tau_\xi\varphi^\star \colon \xi \in \mathbb{R}\}.$$

(ii) For every $\varepsilon > 0$,

$$\inf\{\mathcal{G}(\varphi) \colon \varphi \in \mathcal{F}, \operatorname{dist}(\varphi,\mathcal{M}) \geq \varepsilon\} > \chi.$$

The proof of Proposition 1.1 is given in Section 2.

The second proposition formulates a quantitative statement about the rates of convergence in Theorem 1(i) and (ii), that is, for the asymptotics of the partition sum and for the Q_n–probability of $\{\operatorname{dist}(\overline{\ell}_n,\mathcal{M}) \geq \varepsilon\}$. In particular, the next proposition implies that all accumulation points of $Q_n(\overline{\ell}_n \in \cdot\,)$ are concentrated on $\mathcal{M}$.

Proposition 1.2 (Partition sum and rate of convergence) *There exists $R_0 > 0$ such that,*

(i) for some $N_0 > 0$,

$$e^{-\chi\alpha_n \log n} e^{-R^\star\alpha_n \log\log n} \leq Z_n \leq e^{-\chi\alpha_n \log n} e^{R_0\alpha_n \log\log n}, \qquad n \geq N_0,$$

(ii) for any $\varepsilon > 0$, there is a $C_\varepsilon > 0$ and $N_1 = N_1(\varepsilon)$ such that

$$E\left(e^{-H_n}\mathbb{1}\{\mathrm{dist}(\overline{\ell}_n, \mathcal{M}) \geq \varepsilon\}\right) \leq e^{-(\chi+C_\varepsilon)\alpha_n \log n}\, e^{R_0\alpha_n \log\log n}, \qquad n \geq N_1. \tag{1.2}$$

From Propositions 1.1 and 1.2, the assertions (i) and (ii) of Theorem 1 follow immediately.

The third proposition is the main step in the proof of convergence of $|\operatorname{supp}(\overline{\ell}_n)|$. We call the number $\#\operatorname{supp}\ell_n = \#\{S_0, \dots, S_n\}$ the *range* of the polymer.

Proposition 1.3 (Exponential bound for the range) *There is a constant $C > 0$ such that, for all $\delta > 0$ and all sufficiently large $n \in \mathbb{N}$,*

$$Q_n(\#\operatorname{supp}\ell_n > 2(R^\star + \delta)\alpha_n) \leq e^{-C\delta\alpha_n}. \tag{1.3}$$

Assertion (iii) of Theorem 1 follows from Proposition 1.3 together with Proposition 1.2. Indeed, from Proposition 1.2, one knows that $\frac{1}{\alpha_n}\#\operatorname{supp}\ell_n$ is bigger than $2(R^\star - \delta)$ with Q_n-probability tending to one, and Proposition 1.3 rules out that $\frac{1}{\alpha_n}\#\operatorname{supp}\ell_n$ is larger than $2(R^\star + \delta)$.

2 Proof of Proposition 1.1: The variational problem

In this section we prove Proposition 1.1. We compute the minimizer of the variational problem (0.13) and prove that it is unique up to translations. Finally, we show that for a function $\psi \in \mathcal{F}$ in order to have a small value $\mathcal{G}(\psi)$ it has to be close to the set $\mathcal{M}$ in $\|\cdot\|$-sense.

We divide the proof in several steps. We need to introduce, for $R > 0$, the set $\mathcal{F}_R = \{\psi \in \mathcal{F}\colon \operatorname{supp}(\psi) = [-R, R]\}$ and the function $\varphi_R \in \mathcal{F}_R$ given by $\varphi_R(x) = \frac{3}{4R}\left(1 - (x/R)^2\right)_+$. Recall that we suppress the dependence on β from the notation. First we show that φ_R is the unique minimizer of $\mathcal{G}$ on $\mathcal{F}_R$:

STEP 1 *For any $\psi \in \mathcal{F}_R$,*

$$\mathcal{G}(\psi) - \mathcal{G}(\varphi_R) \geq \frac{\beta}{16R^3}\,\|\psi - \varphi_R\|_1^2. \tag{2.1}$$

Proof. Note that the second derivative φ_R'' is constant on $(-R, R)$ and thus $\int_{-R}^R \varphi_R''(x)(\varphi_R(x) - \psi(x))\,dx = 0$. Furthermore, $\varphi_R'(\varphi_R - \psi)\big|_{-R}^R = 0$. Hence, partial integration yields

$$\begin{aligned} \mathcal{G}(\psi) - \mathcal{G}(\varphi_R) &= \frac{\beta}{2}\int_{-R}^R \left((\psi')^2(x) - (\varphi_R')^2(x)\right) dx = \frac{\beta}{2}\int_{-R}^R (\psi'(x) - \varphi_R'(x))^2\,dx \\ &\geq \frac{\beta}{4R}\left(\int_{-R}^R \left|\psi'(x) - \varphi_R'(x)\right| dx\right)^2 \geq \frac{\beta}{16R^3}\|\psi - \varphi_R\|_1^2. \end{aligned} \tag{2.2}$$

□

Next we show that $\mathcal{G}(\varphi_R)$ is minimal precisely in $R = R^\star$:

STEP 2 *The map $(0, \infty) \to \mathbb{R}$, $R \mapsto \mathcal{G}(\varphi_R)$ is minimal precisely in $R = R^\star = (\frac{9}{2}\beta)^{1/4}$. The minimal value is $\chi = \mathcal{G}(\varphi_{R^\star}) = \frac{\sqrt[4]{8}}{\sqrt{3}}\,\beta^{1/4}$.*

Proof. We compute

$$\mathcal{G}(\varphi_R) = \frac{1}{4}\cdot 2R + \frac{\beta}{2}\int_{-R}^R \left(\frac{3}{4R}\right)^2\left(\frac{2}{R^2}\right)^2 x^2\,dx = \frac{R}{2} + \frac{9\beta}{8R^6}\cdot\frac{2}{3}R^3 = \frac{R}{2} + \frac{3\beta}{4}R^{-3}.$$

Minimizing yields $R = R^\star = (\frac{9}{2}\beta)^{1/4}$ and $\mathcal{G}(\varphi_{R^\star}) = \frac{\sqrt[4]{8}}{\sqrt{3}}\,\beta^{1/4}$. □

Together with Step 1 this implies that $\varphi^\star = \varphi_{R^\star}$ is the unique (up to shifts) minimizer of (0.13) with connected support. In the next step we show that we do not have to consider functions whose support is not connected.

STEP 3 *Let $\psi_1, \psi_2 \in \mathcal{F}$ with disjoint supports. Then for every $\alpha \in [0,1]$*

$$\mathcal{G}\big(\alpha\psi_1 + (1-\alpha)\psi_2\big) \geq \big(\sqrt{\alpha} + \sqrt{1-\alpha}\big)\chi. \tag{2.3}$$

Proof. We have

$$\begin{aligned} \mathcal{G}\big(\alpha\psi_1 + (1-\alpha)\psi_2\big) &= \alpha^2 \tfrac{\beta}{2}\|\psi_1'\|_2^2 + \tfrac{1}{4}|\operatorname{supp}(\psi_1)| + (1-\alpha)^2\tfrac{\beta}{2}\|\psi_2'\|_2^2 + \tfrac{1}{4}|\operatorname{supp}(\psi_2)| \\ &\geq \mathcal{G}_{\alpha^2\beta}(\varphi^\star) + \mathcal{G}_{(1-\alpha)^2\beta}(\varphi^\star) \\ &= \frac{\sqrt[4]{8}}{\sqrt{3}}\beta^{1/4}\big(\sqrt{\alpha} + \sqrt{1-\alpha}\big). \end{aligned} \tag{2.4}$$

□

Steps 1–3 prove assertion (i) of Proposition 1.1.

In the next step we estimate the norm $\|\cdot\|$ in terms of the norm $\|\cdot\|_1$.

STEP 4 *Assume that $\psi \in L^1(\mathbb{R})$ is continuous and almost everywhere differentiable. Then*

$$\|\psi\| \leq \|\psi\|_1 + \|\psi\|_1^{1/3} \cdot \|\psi'\|_2^{2/3}. \tag{2.5}$$

Proof. If $\|\psi'\|_2 = \infty$ then (2.5) is trivially true. Hence we may assume that $\|\psi'\|_2 < \infty$. In this case (since $\|\psi\|_1 < \infty$) $\lim_{|x|\to\infty} \psi(x) = 0$. Thus

$$\psi(x)^2 = 2\int_{-\infty}^{x} \psi'(t)\psi(t)\,dt = -2\int_{x}^{\infty} \psi'(t)\psi(t)\,dt = \int_{-\infty}^{\infty} \psi'(t)\psi(t)\operatorname{sign}(t-x)\,dt \tag{2.6}$$

Applying the Cauchy-Schwarz inequality gives

$$\|\psi\|_\infty^2 \leq \|\psi\|_2 \cdot \|\psi'\|_2 \leq \|\psi\|_\infty^{1/2}\|\psi\|_1^{1/2} \cdot \|\psi'\|_2, \tag{2.7}$$

which implies

$$\|\psi\|_\infty \leq \|\psi\|_1^{1/3} \cdot \|\psi'\|_2^{2/3}. \tag{2.8}$$

From this, (2.5) follows.

□

We come to the final statement of this section. The assertion (ii) of Proposition 1.1 is equivalent to the assertion of the following step.

STEP 5 *Assume that $(\psi_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{F}$ with $\lim_{n\to\infty}\mathcal{G}(\psi_n) = \chi$. Then (up to shifts) ψ_n approaches $\varphi^\star$:*

$$\lim_{n\to\infty} \operatorname{dist}(\psi_n, \mathcal{M}) = 0.$$

Proof. By Step 3 we may assume that $\mathrm{supp}(\psi_n)$ is connected for every $n \in \mathbb{N}$. We may also assume that the supports are centered:

$$\mathrm{supp}(\psi_n) = [-R_n, R_n] \qquad \text{for some } R_n > 0, \quad n \in \mathbb{N}. \tag{2.9}$$

Obviously $R := \sup_{n\in\mathbb{N}} R_n < \infty$. Hence, by Step 1, we have $\lim_{n\to\infty} \|\psi_n - \varphi_{R_n}\|_1 = 0$. From Step 1 and Step 2 it is clear that $\lim_{n\to\infty} R_n = R^\star$, thus $\lim_{n\to\infty} \|\varphi_{R_n} - \varphi_{R^\star}\| = 0$.

It remains to show that $\lim_{n\to\infty} \|\psi_n - \varphi_{R_n}\| = 0$. Note that $\|\psi_n'\|_2$ is bounded since $\mathcal{G}(\varphi) \geq \frac{\beta}{2}\|\varphi'\|_2^2$ for any $\varphi \in \mathcal{F}$. Use (2.5) to see that

$$\|\psi_n - \varphi_{R_n}\| \leq \|\psi_n - \varphi_{R_n}\|_1 + \text{const.}\, \|\psi_n - \varphi_{R_n}\|_1^{1/3},$$

which vanishes as $n \to \infty$. This finishes the proof. $\square$

The proof of Proposition 1.1 is now complete.

3 Preparations

In this section, we prepare for the proofs of Propositions 1.2 and 1.3. In Section 3.1 we describe the distribution of the local times by means of Knight's Theorem, and in Section 3.2 we prove a lower bound for the partition function by investigating the largest Q_n-atom for the local times.

3.1 Using Knight's theorem

We give a characterization of the joint distribution of the walker's local times $\ell_n = (\ell_n(x))_{x\in\mathbb{Z}}$ and the endpoint S_n for fixed $n \in \mathbb{N}$. Our technique is based on Knight's Theorem which is the discrete version of the so-called Ray-Knight Theorem for the local times of Brownian motion.

Our Hamiltonian is not only a function of the walker's local times, but can also be written in terms of the walker's number of upsteps by time n, given by

$$m_n(x) = \sum_{i=1}^{n} 1\!\!1\{S_{i-1} = x, S_i = x+1\}, \qquad x \in \mathbb{Z}, n \in \mathbb{N}. \tag{3.1}$$

Indeed, we have, for $S_n > 0$,

$$\ell_n(x) = m_n(x) + m_n(x-1) - 1\!\!1_{[1,S_n-1]}(x), \qquad x \in \mathbb{Z}, n \in \mathbb{N}, \tag{3.2}$$

and an analogous formula holds for $S_n \leq 0$. Note that, given S_n, the sequence $m_n = (m_n(x))_{x\in\mathbb{Z}}$ of upsteps is uniquely determined by the sequence $\ell_n = (\ell_n(x))_{x\in\mathbb{Z}}$ of local times.

The nice thing about the description in terms of the upsteps rather than local times is that the sequence m_n has an accessible distribution. Up to our best knowledge, the following observation first entered the literature in [Kn63] and has been rediscovered several times since then.

Recall that we have defined

$$P(i,j) = \binom{i+j-2}{i-1}\left(\frac{1}{2}\right)^{i+j-1}, \qquad i,j \in \mathbb{N}.$$

Roughly speaking, Knight's theorem states that, given S_n, the sequence $(m_n(S_n - x))_{x=0,\dots,S_n}$ is a Markov chain on $\mathbb{N}$ with transition kernel $P(\,\cdot\,,\,\cdot\,)$, and that $(m_n(S_n + x))_{x\in\mathbb{N}_0}$ and $(m_n(-x))_{x\in\mathbb{N}_0}$ are Markov chains on $\mathbb{N}_0$ with transition kernel $P^*(\,\cdot\,,\,\cdot\,)$ given by

$$P^*(i,j) = P(i,j+1)1\!\!1_{\mathbb{N}}(i) + 1\!\!1_{\{(0,0)\}}(i,j), \qquad i,j \in \mathbb{N}_0.$$

Note that $P(i,\cdot)$ is the distribution of 1 + the sum of i independent variables which are geometrically distributed on $\mathbb{N}_0$ with parameter $\frac{1}{2}$. Thus, $P(\cdot,\cdot)$ is the transition kernel of a critical branching process with geometrical offspring distribution and with one immigrant per time unit. Furthermore, $P^*(\cdot,\cdot)$ is the transition kernel of a critical branching process on $\mathbb{N}_0$ with geometrical offspring distribution on $\mathbb{N}_0$ with parameter $1/2$, that is, with zero as an absorbing boundary.

Note that, given S_n, the sequence m_n is, with probability one, a random member of the set $I_n(S_n)$ where

$$I_n(s) = \Big\{\mathbf{i} = (i(x))_{x\in\mathbb{Z}} \in \mathbb{N}_0^{\mathbb{Z}} : \sum_{x\in\mathbb{Z}} i(x) = \frac{1}{2}(n+s), \mathrm{supp}(\mathbf{i}) \supset [0, s-1] \text{ connected}\Big\}. \tag{3.3}$$

The distribution of $m_n = (m_n(x))_{x\in\mathbb{Z}}$ is given in the following lemma. For the most concise formulation, we state the lemma only for paths ending with an upstep. Given $s \in \mathbb{N}$ and $x \in \mathbb{Z}$, we define the matrix P_x^s by

$$P_x^s(i,j) = \begin{cases} P^*(i,j) & \text{if } x \geq s, \\ P(i,j) & \text{if } 0 \leq x < s, \\ P^*(j,i) & \text{if } x < 0. \end{cases} \tag{3.4}$$

Lemma 3.1 (Distribution of the local times) *For any $n \in \mathbb{N}$, any $s \in \mathbb{N}$ and any $\mathbf{i} = (i_x)_{x\in\mathbb{Z}} \in I_n(s)$,*

$$P(S_{n-1} = s-1, S_n = s, m_n = \mathbf{i}) = \prod_{x\in\mathbb{Z}} P_x^s(i_{x-1}, i_x). \tag{3.5}$$

Proof. In order to describe the distribution of m_n given the event $\{S_{n-1} = s-1, S_n = s\}$, we need one Markov chain $(m(x))_{x\in\mathbb{N}_0}$ on $\mathbb{N}_0$ having transition kernel $P(\cdot,\cdot)$ and two Markov chains $(m_1^*(x))_{x\in\mathbb{N}_0}$ and $(m_2^*(x))_{x\in\mathbb{N}_0}$ on $\mathbb{N}_0$ having the transition kernel $P^*(\cdot,\cdot)$. These three chains are defined on one probability space and are assumed to be independent. Now condition on the event

$$\Big\{m_1^*(0) = m(s), m_2^*(0) = m(0), \sum_{x\in\mathbb{N}_0} (m_1^*(x) + m_2^*(x)) + \sum_{x=1}^{s} m(x) = (n+s)/2\Big\}.$$

Then m_n coincides in distribution with $\tilde{m}_n$, defined by

$$\tilde{m}_n(x) = \begin{cases} m_1^*(x-s) & \text{if } x \geq s, \\ m(x) - 1 & \text{if } 0 \leq x < s, \\ m_2^*(-x) & \text{if } x < 0. \end{cases} \tag{3.6}$$

(For the details we refer to [HHK97].) From these facts we obtain (3.5). □

We will be concerned with products of $P(i,j)$'s with large i and j.

Lemma 3.2 (Asymptotics of the transition kernel)

(i) As $i,j \to \infty$, provided that $(i-j)/(i+j) \to 0$,

$$P(i,j) = \frac{\exp\Big\{-\frac{(i-j)^2}{2(i+j)}\Big\}}{\sqrt{2\pi(i+j)}} \left(1 + \mathcal{O}\left(\frac{1}{i} + \frac{1}{j}\right) + \mathcal{O}\left(\frac{(i-j)^3}{(i+j)^2}\right)\right). \tag{3.7}$$

(ii) For any $i, j \in \mathbb{N}$,

$$\max\{P(i,j), P^*(i,j)\} \le \begin{cases} \frac{1}{2}, & \text{if } i = j = 1, \\ \frac{1}{2\sqrt{i+j}}, & \text{otherwise.} \end{cases} \tag{3.8}$$

Proof. *(i).* It is sufficient to derive the assertion for $P(i+1, j+1)$ instead of $P(i,j)$ since

$$P(i,j) = P(i+1, j+1)\left(1 + \mathcal{O}\left(\left(\tfrac{i-j}{i+j}\right)^2\right) + \mathcal{O}\left(\tfrac{1}{i} + \tfrac{1}{j}\right)\right). \tag{3.9}$$

Use Stirling's formula to see that, as $i, j \to \infty$,

$$P(i+1, j+1) = \frac{\exp\left\{-(i+j)\mathcal{S}\left(\frac{i}{i+j}\right)\right\}}{\sqrt{2\pi(i+j)}}\left(1 - \left(\tfrac{i-j}{i+j}\right)^2\right)^{-1/2}\left(1 + \mathcal{O}\left(\tfrac{1}{i} + \tfrac{1}{j}\right)\right)$$

where $\mathcal{S}(p) = p\log p + (1-p)\log(1-p) + \log 2$ denotes (in the jargon of large deviations) the coin-tossing rate function. Using a Taylor expansion we see that

$$2x^2 \le \mathcal{S}(\tfrac{1}{2} + x) \le 2x^2(1 + 4x^2), \qquad x \in (-\tfrac{1}{2}, \tfrac{1}{2}).$$

Applying this with $x = \frac{1}{2}\frac{i-j}{i+j}$ completes the proof.

(ii). Since the right hand side of (3.8) is decreasing in $i+j$, it suffices to show the inequality for $P(i,j)$ only. One checks by direct computation that $\max_{i+j=k} P(i,j) = \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{3}{8}, \frac{3}{16}$, for $k = 2, 3, 4, 5, 6$, thus (3.8) holds for $i + j \le 6$.

Now use Stirling's formula

$$m! = \sqrt{2\pi m}\,(m/e)^m e^{\rho(m)}, \qquad \text{where } (12m+1)^{-1} < \rho(m) < (12m)^{-1}.$$

In particular, ρ is strictly decreasing, thus for $m \ge k \ge 0$, we have $\rho(m) - \rho(k) - \rho(m-k) \le 0$. It follows that

$$\begin{aligned} \binom{m}{k} &\le \frac{1}{\sqrt{2\pi}}\sqrt{\frac{m}{k(m-k)}}\left(\frac{k}{m}\right)^{-k}\left(1 - \frac{k}{m}\right)^{k-m} \\ &\le (2\pi m)^{-1/2}\sup_{x\in[0,1]}\left(x(1-x)\right)^{-1/2}\left(x^x(1-x)^{1-x}\right)^{-m} \\ &= \sqrt{1/\pi}\, m^{-1/2} 2^m, \end{aligned} \tag{3.10}$$

since the supremum is attained for $x = 1/2$. Substituting this estimate into the definition of P yields for $i, j \in \mathbb{N}$ that $P(i,j) \le 1/\sqrt{2\pi(i+j-2)}$. Now use the fact that $2\pi(i+j-2) \ge 4(i+j)$ if $i+j \ge 6$. □

The following corollary gives an upper bound for the contribution coming from all the paths having some fixed local time vector. It will be an indispensable tool in Sections 4 and 5.

For $\delta > 0$, define the functional $\mathcal{G}^\delta \colon \mathcal{F} \to \mathbb{R}$ by

$$\mathcal{G}^\delta(\varphi) = \frac{\beta}{2}\|\varphi'\|_2^2 + \frac{1}{4}|\{\varphi > \delta\}|, \tag{3.11}$$

where we write $\{\varphi > \delta\} = \{t \in \mathbb{R} \colon \varphi(t) > \delta\}$ for short, and $|\cdot|$ denotes the Lebesgue measure.

We introduce the discrete set

$$\mathcal{F}_n = \{\varphi \in \mathcal{F} \colon P(\overline{\ell}_n = \varphi) > 0\}, \qquad n \in \mathbb{N}, \tag{3.12}$$

which is the support of the distribution of $\overline{\ell}_n$ under the simple random walk law P. Recall (3.4).

Corollary 3.3 (A basic upper bound) *For any $n, s \in \mathbb{N}$, every $\varphi \in \mathcal{F}_n$ and any $\delta > 0$,*

$$E\big(e^{-H_n}1\!\!1\{\overline{\ell}_n = \varphi, S_n = s\}\big) \le e^{-\mathcal{G}^\delta(\varphi)\alpha_n \log n} 2n^2 e^{4\sqrt{\beta H_n(\varphi)}} e^{-|\{\varphi>\delta\}|\alpha_n \log \delta} \prod_{x\colon i_x + i_{x-1} \le \delta \frac{n}{\alpha_n}} P_x^s(i_{x-1}, i_x). \qquad (3.13)$$

where

$$H_n(\varphi) = \frac{\beta}{2}\|\varphi'\|_2^2 \alpha_n \log n. \qquad (3.14)$$

Proof. Let us introduce some more notation. In the following, we shall use the multi–index $\mathbf{i} = (i_x)_{x\in\mathbb{Z}} \in I_n(s)$ for some generic upstep vector, we define a generic local time vector $\mathbf{j} = (j_x)_{x\in\mathbb{Z}}$ by $j_x = i_x + i_{x-1} - 1\!\!1\{x \in [1, s-1]\}$ and the scaled linear interpolation $\varphi \in \mathcal{F}$ of $\mathbf{j}$ by $\varphi(\frac{x}{\alpha_n}) = \frac{\alpha_n}{n} j_x$ for $x \in \mathbb{Z}$. Note that

$$m_n = \mathbf{i} \qquad \Longleftrightarrow \qquad \ell_n = \mathbf{j} + 1\!\!1_0 \qquad \Longleftrightarrow \qquad \overline{\ell}_n = \varphi. \qquad (3.15)$$

In terms of this notation, we have that

$$H_n(\varphi) = \frac{\beta}{2}\sum_{x\in\mathbb{Z}}[i_x - i_{x-2}]^2 = \frac{\beta}{2}\sum_{x\in\mathbb{Z}}[j_x - j_{x-1}]^2 = \frac{\beta}{2}\|\varphi'\|_2^2 \alpha_n \log n \qquad (3.16)$$

(recall (0.22)), such that the Hamiltonian H_n is equal to $H_n(\overline{\ell}_n)$. In an abuse of notation, we sometimes shall also write $H_n(\mathbf{i})$ or $H_n(\mathbf{j})$ instead of $H_n(\varphi)$ if no confusion can arise.

Let us first handle paths that end with an upstep. Lemma 3.1 yields

$$\begin{aligned} E\big(e^{-H_n}1\!\!1\{\overline{\ell}_n = \varphi, S_{n-1} = s-1, S_n = s\}\big) &= e^{-H_n(\varphi)} P\big(\overline{\ell}_n = \varphi, S_{n-1} = s-1, S_n = s\big) \\ &= e^{-\frac{\beta}{2}\|\varphi'\|_2^2\alpha_n \log n} \prod_{x\in\mathbb{Z}} P_x^s(i_{x-1}, i_x). \end{aligned} \qquad (3.17)$$

Now extract the product over those x with $i_{x-1} + i_x > \delta\frac{n}{\alpha_n}$, which are $\alpha_n|\{\varphi > \delta\}|$ factors. For every such x, use Lemma 3.2(ii) to estimate $\max\{P(i_{x-1}, i_x), P^*(i_{x-1}, i_x)\} \le \frac{1}{2}j_x^{-1/2} \le (\delta n/\alpha_n)^{-1/2}$. Now use that $\log\frac{n}{\alpha_n} = \frac{1}{2}\log n + \frac{1}{4}\log\log n \ge \frac{1}{2}\log n$ and summarize. This shows that the left hand side of (3.17) is not smaller than the right hand side of (3.13) without the factor $2n^2 e^{3\sqrt{\beta H_n(\varphi)}}$.

In order to handle the paths that end with a downstep at time n, note that a flip of the last step does not change the path's probability, and it corresponds to a switch from s to $s+2$, from $\mathbf{i}$ to $\mathbf{i} + 1\!\!1_{s+1}$ respectively from $\mathbf{j}$ to $\mathbf{j} - 1\!\!1_s + 1\!\!1_{s+2}$:

$$P\big(\overline{\ell}_n = \varphi, S_{n-1} = s+1, S_n = s\big) = P\big(\ell_n = \mathbf{j} - 1\!\!1_s + 1\!\!1_{s+2}, S_{n-1} = s+1, S_n = s+2\big).$$

Now apply the above to these new parameters and note that

$$\frac{P_{s+2}^{s+2}(i_{s+1}+1, i_{s+2})P_{s+1}^{s+2}(i_s, i_{s+1}+1)}{P_{s+2}^{s}(i_{s+1}, i_{s+2})P_{s+1}^{s}(i_s, i_{s+1})} = \frac{(i_{s+2}+i_{s+1})(i_{s+1}+i_s)}{4i_{s+1}^2} \le n^2, \qquad (3.18)$$

and

$$|H_n(\mathbf{i} + 1\!\!1_{s+1}) - H_n(\mathbf{i})| = |\beta + \beta(2i_{s+1} - i_{s-1} - i_{s+3})| \le 4\sqrt{\beta H_n(\varphi)}. \qquad (3.19)$$

Substituting these bounds proves the claim.

$\square$

3.2 The largest Q_n-atom for the local times

In the following lemma, we apply the results of Section 3.1 for deriving a crucial lower bound. It shows that in our model the entropy effect arises only on the level of the path but not on the level of the local times. In particular, the lower bound in Proposition 1.2(i) follows from Lemma 3.4.

Lemma 3.4 (The lower bound) *For sufficiently large $n \in \mathbb{N}$,*

$$\max_{\varphi \in \mathcal{F}_n} E\left(e^{-H_n} \mathbb{1}\{\overline{\ell}_n = \varphi\}\right) \geq e^{-\chi \alpha_n \log n} e^{-\frac{R^\star}{4} \alpha_n \log \log n}. \tag{3.20}$$

Proof. We prove the statement for n even. The proof for n odd is similar.

We pick some approximation φ of $\varphi^\star$ as follows. Choose some $\phi \in \mathcal{F}_{n/2}$ such that $\operatorname{supp} \phi \subset [-R^\star, R^\star]$ and such that $i_x = \frac{1}{2} \frac{n}{\alpha_n} \phi(\frac{x}{\alpha_n}) \in \mathbb{N}_0$ and such that $-1 \leq u_x \leq 1$ for any $x \in \mathbb{Z}$, where

$$u_x = \frac{1}{2} \frac{n}{\alpha_n} \left(\phi\left(\tfrac{x}{\alpha_n}\right) - \varphi^\star\left(\tfrac{x}{\alpha_n}\right)\right). \tag{3.21}$$

Note that $\mathbf{i} = (i_x)_{x \in \mathbb{Z}}$ lies in $I_n(0)$. The above ϕ corresponds to a $\varphi \in \mathcal{F}_n$ via $\varphi(t) = \phi(t - 1/\alpha_n) + \phi(t)$. That is, $\varphi(x/\alpha_n) = (n/\alpha_n) j_x$, where $\mathbf{j}$ is given by $j_x = i_x + i_{x-1}$ (recall (3.15) and use that $s = 0$). This construction makes sure that both φ and its upstep vector function ϕ are close to $\varphi^\star$ respectively $\frac{1}{2}\varphi^\star$.

We claim that, on the event $\{\overline{\ell}_n = \varphi\}$, we have

$$\begin{aligned}
&(i) \quad H_n = \frac{\beta}{2} \|(\varphi^\star)'\|_2^2 \, \alpha_n \log n \left(1 + o\left(\frac{1}{\log n}\right)\right), \\
&(ii) \quad \sum_{x \in \mathbb{Z}} \frac{(i_x - i_{x-1})^2}{j_x} \leq o(\alpha_n), \\
&(iii) \quad \prod_x j_x^{-\frac{1}{2}} \geq e^{-R^\star \alpha_n \log \frac{n}{\alpha_n}} = e^{-\frac{R^\star}{2} \alpha_n \log n} e^{-\frac{R^\star}{4} \alpha_n \log \log n}.
\end{aligned}$$

To see that (i) holds, note that

$$H_n = \frac{\beta}{2} \sum_{x \in \mathbb{Z}} [i_x - i_{x-2}]^2 = \frac{\beta}{2} \sum_{x \in \mathbb{Z}} \left[\frac{n}{2\alpha_n}\left(\varphi^\star\left(\frac{x}{\alpha_n}\right) - \varphi^\star\left(\frac{x-2}{\alpha_n}\right)\right) + u_x - u_{x-2}\right]^2. \tag{3.22}$$

Writing out the square gives three terms, the first of which is

$$\left(\frac{n}{2\alpha_n}\right)^2 \sum_{x \in \mathbb{Z}} \left[\varphi^\star\left(\frac{x}{\alpha_n}\right) - \varphi^\star\left(\frac{x-2}{\alpha_n}\right)\right]^2 = \int_{\mathbb{R}} (\varphi^\star)'(t)^2 \, dt \, \alpha_n \log n \left(1 + o\left(\frac{1}{\log n}\right)\right). \tag{3.23}$$

The third term is bounded in absolute value by $|\sum_x [u_x - u_{x-2}]^2| \leq 8 R^\star \alpha_n$. The cross term is bounded from above using partial summation by

$$2 \frac{n}{\alpha_n} \sum_{x \in \mathbb{Z}} |u_x| \left|\varphi^\star\left(\frac{x+2}{\alpha_n}\right) - 2\varphi^\star\left(\frac{x}{\alpha_n}\right) + \varphi^\star\left(\frac{x+2}{\alpha_n}\right)\right| \leq 8 \|(\varphi^\star)''\|_1 \frac{n}{\alpha_n^2} = o(\alpha_n), \tag{3.24}$$

where we used that

$$\left|\varphi^\star\left(\frac{x+2}{\alpha_n}\right) - 2\varphi^\star\left(\frac{x}{\alpha_n}\right) + \varphi^\star\left(\frac{x+2}{\alpha_n}\right)\right| \leq \frac{2}{\alpha_n} \int_{(x-2)/\alpha_n}^{(x+2)/\alpha_n} |(\varphi^\star)''(t)| \, dt.$$

In a similar way, one derives assertions (ii) and (iii). Now use Lemmas 3.1 and 3.2 and the fact that $\chi = \frac{\beta}{2} \|(\varphi^\star)'\|_2^2 + R^\star/2$ to obtain the assertions. □

4 Proof of Proposition 1.2

In this section we show the upper bounds of Proposition 1.2. Recall that the lower bound in (i) was already established in Lemma 3.4.

In Section 4.1 we show that the contribution to e^{-H_n} that comes from paths satisfying $|\operatorname{supp}(\overline{\ell}_n)| > R\log n$ or $|\{t\colon \overline{\ell}_n(t) > (\log n)^{-2}\}| > R$ is extremely small if R is large enough. This is done by a precise analysis of the probability of these events under the free walk measure, using the Knight description of the walker's local times.

In Section 4.2, we take advantage of having established the additional constraints $|\operatorname{supp}(\overline{\ell}_n)| \leq R\log n$ and $|\{t\colon \overline{\ell}_n(t) > (\log n)^{-2}\}| \leq R$ for some $R > 0$ in order to show that also the contribution from paths satisfying $\operatorname{dist}(\overline{\ell}_n, \mathcal{M}) > \varepsilon$ is much smaller than the partition sum Z_n, which is the assertion of Proposition 1.2.

4.1 Bounding the range of the polymer

After two preliminary lemmas (Lemma 4.1 and 4.2) supplying us with a volume estimate and a bound on binomial coefficients, we come to the first main statement: In Lemma 4.3 we give exponential bounds for the Q_n–probability of a too large support of the polymer. This lemma will be pivotal also for the proof of Proposition 1.3 in Section 5.

For any function $f\colon \mathbb{Z} to \mathbb{N}_0$ and $K \in \mathbb{R}$, we abbreviate $\{f > K\} = \{x \in \mathbb{Z}\colon f(x) > K\}$ and denote the cardinality of this set by $\#\{f > K\}$. Note that $\operatorname{supp}\ell_n = \{S_0, \dots, S_n\}$.

Lemma 4.1 (Volume estimate) *For all sufficiently large $R_1, R_2, M \in \mathbb{N}$ such that $R_1 > M$, and for either $K = 1$ or $K \geq 3$ with $K < M$,*

$$\sup_{n\in\mathbb{N}} P\Big(\#\operatorname{supp}\ell_n = R_1, \#\{\ell_n \geq K\} = R_2, H_n \leq \frac{\beta}{2}M\Big) \leq R_1 2^{-R_1} K^{-R_2/2}\binom{R_1}{M} 8^M. \tag{4.1}$$

Proof. Let

$$B^{M,K}_{R_1,R_2} = \Big\{\mathbf{i} = (i_x)_{x=1}^{R_1} \in \mathbb{N}^{R_1} : \sum_{x=1}^{R_1}(i_x - i_{x-2})^2 \leq M, \#\{x\colon i_{x-1} + i_x \geq K\} = R_2\Big\}, \tag{4.2}$$

where we put $i_{R_1+1} = 0$. Note that $B^{M,1}_{R_1,R_2}$ is void unless $R_1 \geq R_2$. Bound the left hand side of (4.1) from above by

$$R_1 \sum_{\mathbf{i}\in B^{M,K}_{R_1,R_2}} \prod_{x=1}^{R_1} \max\{P(i_{x-1}, i_x), P^*(i_{x-1}, i_x)\}. \tag{4.3}$$

For either $K = 1$ and all x or for large K and x such that $i_{x-1} + i_x < K$, we use (see Lemma 3.2(ii)) the rough estimate

$$\max\{P(i_{x-1}, i_x), P^*(i_{x-1}, i_x)\} \leq \frac{1}{2},$$

while for x such that $i_x + i_{x-1} \geq K$, we use that, for $K \geq 3$,

$$\max\{P(i_{x-1}, i_x), P^*(i_{x-1}, i_x)\} \leq \frac{1}{2\sqrt{K}}.$$

Hence,

$$\text{l.h.s. of (4.1)} \leq R_1\Big(\frac{1}{2}\Big)^{R_1-R_2}\Big(\frac{1}{2\sqrt{K}}\Big)^{R_2}\#B^{M,K}_{R_1,R_2} = R_1 2^{-R_1} K^{-\frac{1}{2}R_2}\#B^{M,K}_{R_1,R_2}. \tag{4.4}$$

The first three factors give the first three factors in the right hand side of (4.1). Hence, we are left to bound $\#B^{M,K}_{R_1,R_2}$. This is done as follows.

$$
\begin{aligned}
\#B^{M,K}_{R_1,R_2} &\le \#\Big\{\mathbf{i}=(i_x)_{x=1}^{R_1}\in\mathbb{N}^{R_1}\colon \sum_{x=1}^{R_1}(i_x-i_{x-2})^2\le M\Big\}\\
&\le \#\Big\{\mathbf{d}=(d_x)_{x=1}^{R_1}\in\mathbb{Z}^{R_1}\colon \sum_{x=1}^{R_1}d_x^2\le M\Big\}\\
&\le \binom{R_1}{M}2^M\#\Big\{\mathbf{d}=(d_x)_{x=1}^{M}\in\mathbb{N}^{M}\colon \sum_{x=1}^{M}d_x\le M\Big\}\\
&\le \binom{R_1}{M}2^M\sum_{k=0}^{M}\#\{\mathbf{d}\in\mathbb{N}^M\colon \|\mathbf{d}\|_1=k\}\\
&= \binom{R_1}{M}2^M\sum_{k=0}^{M}\binom{M+k-1}{M}.
\end{aligned}
\tag{4.5}
$$

The proof now follows from the bound $\binom{M+k-1}{M}\le 2^{M+k-1}$ and performing the sum over k. □

In the proof of the next corollary and later we need the following elementary estimate for binomials.

Lemma 4.2 (Bound for binomial coefficients) *For all $m,k\in\mathbb{N}$ such that $e^2k\le m$,*

$$
\binom{m+k}{m}\le e^{2k\log(m/k)}. \tag{4.6}
$$

Proof. Using Stirling's formula as in (3.10), we see that

$$
\begin{aligned}
\binom{m+k}{m} &\le \frac{1}{\sqrt{2\pi}}\sqrt{\frac{m+k}{mk}}\,\frac{(m+k)^{(m+k)}}{k^km^m}\\
&\le \frac{(m+k)^{(m+k)}}{k^km^m}=\Big(1+\frac{k}{m}\Big)^{m+k}\Big(\frac{m}{k}\Big)^k\\
&= e^{2k\log(m/k)}\exp\Big((m+k)\log\Big(1+\frac{k}{m}\Big)-k\log\frac{m}{k}\Big).
\end{aligned}
\tag{4.7}
$$

By assumption $m\ge e^2k$, so that by the estimate $\log(1+x)\le x$ it follows that

$$
(m+k)\log\Big(1+\frac{k}{m}\Big)-k\log\frac{m}{k}\le\frac{m+k}{m}k-2k\le 0. \tag{4.8}
$$

□

Based on Lemmas 4.1 and 4.2, we can now prove relatively easily a weaker version of Proposition 1.3. In particular, we rule out that the support of $\overline{\ell}_n$ is longer than $\log n$ times a large constant.

Lemma 4.3 (Exponential bounds on the range of the polymer) *There exist $C,R_0,N_0>0$ such that, for all $n\ge N_0$,*

(i) for all $L\ge R_0\alpha_n\log n$,

$$
Q_n\big(\#\operatorname{supp}\ell_n>2L\big)\le e^{-CL}, \tag{4.9}
$$

(ii) for all $R \geq R_0$,

$$Q_n\big(\#\{\ell_n > n^{\frac{1}{8}}\} > R\alpha_n\big) \leq e^{-CR\alpha_n \log n}. \tag{4.10}$$

Proof. *(i).* Estimate, for all $c \in (0,1)$,

$$Q_n\big(\#\operatorname{supp}\ell_n > 2L\big) \leq \frac{e^{-\beta cL/2}}{Z_n} + \frac{1}{Z_n} P\Big(\#\operatorname{supp}\ell_n > 2L, H_n \leq \frac{\beta cL}{2}\Big). \tag{4.11}$$

In order to further estimate the second term, we apply Lemma 4.1 for $M = \lfloor cL \rfloor$ for some small $c > 0$, $K = 1$ and $R_1 = R_2$ and sum over $R_1 \geq 2L$, where we determine c in the course of the proof. This yields

$$\begin{aligned} P\Big(\#\operatorname{supp}\ell_n > 2L, H_n \leq \frac{\beta cL}{2}\Big) &\leq 8^{cL} \sum_{R_1 \geq 2L} R_1 2^{-R_1} \binom{R_1}{\lfloor cL \rfloor} \\ &\leq e^{(3\log 2)cL} \sum_{R_1 \geq 2L} R_1 2^{-R_1} e^{2cL \log \frac{R_1 - cL}{cL}} \\ &\leq e^{(3\log 2 - 2\log(2/c))cL} \sum_{R_1 \geq 2L} R_1 e^{(c - \log 2)R_1}, \end{aligned} \tag{4.12}$$

where in the second line we used that L is large, and we applied Lemma 4.2 with $k = \lfloor cL \rfloor$, and $m = R_1 - \lfloor cL \rfloor$. To be able to apply Lemma 4.2, we need that $e^2 k = e^2 \lfloor cL \rfloor \leq m = R_1 - \lfloor cL \rfloor$, which is true when $c \leq 2/(1+e^2)$. Assuming even $c \in (0, 0.01)$ we have $\log 2 - c \geq 2/3$, thus a simple calculation shows that for $L \geq 4$

$$\sum_{R_1 \geq 2L} R_1 e^{(c-\log 2)R_1} \leq 4Le^{(c-\log 2)L}.$$

For this choice of c also $c(3\log 2 - 2\log(2/c)) < (\log 2)/4$ and $c < (\log 2)/4$, thus (4.12) can be continued by

$$\leq 4Le^{-L(\log 2)/4} e^{-L\log 2} \leq e^{-L\log 2}$$

for L large enough.

Hence, using this in (4.11), we see that, for sufficiently small $c \in (0,1)$, the r.h.s. of (4.11) is bounded by $\frac{1}{Z_n} e^{-CL}$ for some $C > 0$. Using Corollary 3.4 and recalling that $L \geq R_0 \alpha_n \log n$, we arrive at the assertion (after possibly changing the value of C).

(ii). This is analogous to the proof of (i), and we point out the differences only. Start with an estimate analogously to (4.11). This time apply Lemma 4.1 for $M = \lfloor cR\alpha_n \log n \rfloor$ and $K = \lfloor n^{1/8} \rfloor$ and sum over $R_1 \geq R_2 \geq R\alpha_n$. Again use Lemma 4.2 and summarize to get the assertion. □

4.2 Using the constraints

In Section 4.1 we have established bounds for the Q_n–probability that the range of the polymer is too large. In Lemma 4.4 we will give an upper bound for the contribution to the partition sum coming from paths with not too large a range. This puts us in the position to finish the proof of Proposition 1.2 at the end of this section.

For the next lemma, recall from Proposition 1.1(ii) that, for any $\varepsilon > 0$, the constant

$$C^{(\varepsilon)} := \inf\big\{\mathcal{G}(\varphi) : \operatorname{dist}(\varphi, \mathcal{M}) \geq \varepsilon\big\} - \chi \tag{4.13}$$

is positive. For $R, M > 0$, introduce the set

$$\mathcal{F}_n(R,M) = \big\{\varphi \in \mathcal{F}_n : |\operatorname{supp}(\varphi)| \leq R\log n, |\{\varphi > (\log n)^{-2}\}| \leq R, \|\varphi'\|_2^2 \leq M\alpha_n \log n\big\}, \tag{4.14}$$

where we recall that we write $\{\varphi > \delta\}$ short for $\{t \in \mathbb{R} \colon \varphi(t) > \delta\}$. We have already shown in Lemma 4.3 that $\lim_{n\to\infty} Q_n(\overline{\ell}_n \in \mathcal{F}_n(R,M)^{\mathrm{c}}) = 0$ if R and M are large.

Lemma 4.4 (The main upper bound) *Fix $M, R > 0$ and $\varepsilon \geq 0$. Then, for all sufficiently large $n \in \mathbb{N}$,*

$$E\big(e^{-H_n} \mathbb{1}\{\mathrm{dist}(\overline{\ell}_n, \mathcal{M}) \geq \varepsilon\} \mathbb{1}\{\overline{\ell}_n \in \mathcal{F}_n(R, M)\}\big) \leq e^{-(\chi + C^{(\varepsilon/3)})\alpha_n \log n}\, e^{8R\alpha_n \log\log n}. \tag{4.15}$$

Proof. Abbreviate $\delta_n = (\log n)^{-2}$ and $\varepsilon_n = \varepsilon - (3 + 4/(3R^\star))R/\log n = \varepsilon + o(1)$, where we let $\varepsilon_n = 0$ for $\varepsilon = 0$. First we show that for any $\varphi \in \mathcal{F}_n(R, M)$ such that $\mathrm{dist}(\varphi, \mathcal{M}) \geq \varepsilon$, we have

$$\mathcal{G}^{\delta_n}(\varphi) \geq (\chi + C^{(\varepsilon_n)})\Big(1 - \frac{R}{\log n}\Big)^2. \tag{4.16}$$

In order to do this, define $\varphi_n(x) = (\varphi(x) - \delta_n)_+$ and note that, since $\|\varphi_n\|_1 \geq 1 - \delta_n |\operatorname{supp}(\varphi_n)| \geq 1 - R\delta_n \log n = 1 - R/\log n$,

$$\mathcal{G}^{\delta_n}(\varphi) \geq \mathcal{G}(\varphi_n) \geq \mathcal{G}\left(\frac{\varphi_n}{\|\varphi_n\|_1}\right)\|\varphi_n\|_1^2 \geq \mathcal{G}\left(\frac{\varphi_n}{\|\varphi_n\|_1}\right)\left(1 - \frac{R}{\log n}\right)^2. \tag{4.17}$$

In order to show that (4.16) holds, it is enough to show that $\mathrm{dist}(\varphi_n/\|\varphi_n\|_1, \mathcal{M}) \geq \varepsilon_n$. For doing this, use that $1 \geq \|\varphi_n\|_1$ and the triangle inequality to get

$$\mathrm{dist}(\varphi_n/\|\varphi_n\|_1, \mathcal{M}) \geq \mathrm{dist}(\varphi_n,\ \|\varphi_n\|_1\mathcal{M}) \geq \mathrm{dist}(\varphi,\ \mathcal{M}) - \|\varphi - \varphi_n\| - (1 - \|\varphi_n\|_1)\|\varphi^\star\|. \tag{4.18}$$

Now recall that $\mathrm{dist}(\varphi, \mathcal{M}) \geq \varepsilon$, that $\|\varphi^\star\| = 1 + \frac{4}{3R^\star}$ and that $1 - \|\varphi_n\|_1 \leq R/\log n$ and observe that $\|\varphi - \varphi_n\| \leq 2R\delta_n \log n$ to conclude that $\mathrm{dist}(\varphi_n/\|\varphi_n\|_1, \mathcal{M}) \geq \varepsilon_n$. Hence (4.16) holds. In particular, we have, for sufficiently large $n \in \mathbb{N}$,

$$\mathcal{G}^{\delta_n}(\varphi) \geq \chi + C^{(\varepsilon/2)}. \tag{4.19}$$

Abbreviate $R_1 = \lfloor R\alpha_n \log n\rfloor$ and put

$$B_R^n = \left\{\mathbf{i} = (i_x)_{x=1}^{R_1} \in \mathbb{N}_0^{R_1} : \sum_{x=1}^{R_1}(i_x - i_{x-2})^2 \leq M\alpha_n \log n, \#\Big\{x : i_{x-1} + i_x > \delta_n \frac{n}{\alpha_n}\Big\} \leq R\alpha_n\right\}.$$

We bound $2n^2 e^{4\sqrt{\beta H_n(\varphi)}} \leq e^{\frac{1}{2}R\alpha_n \log \delta_n}$. Then, by Corollary 3.3 and (4.19), we have, for large n,

$$\text{l.h.s. of (4.15)} \leq e^{-(\chi + C^{(\varepsilon/2)})\alpha_n \log n}\, e^{-\frac{1}{2}R\alpha_n \log \delta_n} \sum_{s=-n}^{n} \sum_{\mathbf{i} \in B_R^n} \prod_{x :\, i_x + i_{x-1} \leq \delta_n \frac{n}{\alpha_n}} P_x^s(i_{x-1}, i_x). \tag{4.20}$$

Note that $\frac{1}{2}R\alpha_n \log \delta_n \geq -3R\alpha_n \log\log n$ for large n. Therefore, to finish the proof of Lemma 4.4, it suffices to show that the sum over $\mathbf{i} \in B_R^n$ on the r.h.s. of (4.20) is not bigger than $\exp\{5R\alpha_n \log\log n\}/(2n+1)$. For $\mathbf{i} \in B_R^n$, define $A(\mathbf{i}) = \big\{x \in \{1, \dots, R_1\} : i_x + i_{x-1} > \delta_n \frac{n}{\alpha_n}\big\}$. Then we can rewrite the sum over $\mathbf{i}$ in (4.20) as

$$\begin{aligned} &\sum_{\substack{A \subset \{1,\dots,R_1\} \\ \#A \leq R\alpha_n}} \sum_{\mathbf{i} \in B_R^n :\, A(\mathbf{i}) = A} \prod_{x \in \{1,\dots,R_1\}\setminus A} P_x^s(i_{x-1}, i_x) \\ &\qquad = \sum_{\substack{A \subset \{1,\dots,R_1\} \\ \#A \leq R\alpha_n}} \sum_{\substack{(d_x)_{x\in A} \\ \sum_{x\in A} d_x^2 \leq M\alpha_n \log n}} \sum_{(i_x)_{x \notin A}} \prod_{x \in \{1,\dots,R_1\}\setminus A} P_x^s(i_{x-1}, i_x), \end{aligned} \tag{4.21}$$

where $d_x = i_x - i_{x-2}$. Now use that P_x^s for any $x \geq 0$ and the transposed of P_x^s for any $x < 0$ are stochastic matrices to perform the sum over $(i_x)_{x \notin A}$, to get that

$$\text{r.h.s. of (4.21)} \leq \#\Big\{(A, (d_x)_{x\in A}) : A \subset \{1, \dots, R_1\}, \#A \leq R\alpha_n, \sum_{x \in A} d_x^2 \leq M\alpha_n \log n\Big\}. \tag{4.22}$$

We proceed as in (4.5) to conclude that

$$\begin{aligned}\text{r.h.s. of (4.21)} &\leq \sum_{1\leq k\leq R\alpha_n} \binom{R_1}{k} \#\{\mathbf{d}\in\mathbb{Z}^k : \|\mathbf{d}\|_2^2 \leq M\alpha_n \log n\} \\ &\leq \sum_{1\leq k\leq R\alpha_n} \binom{R_1}{k} 2^k \#\{\mathbf{d}\in\mathbb{N}_0^k : \|\mathbf{d}\|_1 \leq M\alpha_n \log n\} \\ &\leq \sum_{1\leq k\leq R\alpha_n} \binom{R_1}{k} 2^k \sum_{1\leq l\leq M\alpha_n \log n} \binom{l+k-1}{k}.\end{aligned} \tag{4.23}$$

Recall that $R_1 = R\alpha_n \log n$. We use the fact that $\binom{m}{k}$ and $\binom{m+k}{k}$ are increasing in m to conclude that

$$\begin{aligned}\text{r.h.s. of (4.21)} &\leq \sum_{1\leq k\leq R\alpha_n} \binom{\lfloor R\alpha_n \log n\rfloor}{k} 2^k M\alpha_n \log n \binom{\lfloor M\alpha_n \log n\rfloor + k - 1}{k} \\ &\leq 2^{1+R\alpha_n} M\alpha_n \log n \binom{\lfloor R\alpha_n \log n\rfloor + \lfloor R\alpha_n\rfloor}{\lfloor R\alpha_n\rfloor} \times \binom{\lfloor M\alpha_n \log n\rfloor + \lfloor R\alpha_n\rfloor}{\lfloor R\alpha_n\rfloor}.\end{aligned} \tag{4.24}$$

Use Lemma 4.2 to bound this from above by

$$2^{1+R\alpha_n} M\alpha_n \log n \cdot \exp\left\{2R\alpha_n \left(\log\log n + \log\left(\frac{M}{R}\log n\right)\right)\right\}. \tag{4.25}$$

For large n, this is not bigger than $\exp\{5R\alpha_n \log\log n\}/(2n+1)$. This finishes the proof. □

Proof of Proposition 1.2. The lower bound in (i) follows from Lemma 3.4.

Since $H_n = \frac{\beta}{2}\|\overline{\ell}'_n\|_2^2\,\alpha_n \log n$, one elementarily derives, with the help of Lemma 4.3, if R and M are chosen large enough, that $Q_n(\overline{\ell}_n \in \mathcal{F}_n(R,M)^{\rm c}) \leq e^{-C\alpha_n \log n}$ for sufficiently large $n\in\mathbb{N}$, for some $C = C(R,M) > 0$. Now the upper bound in (i) follows from an application of Lemma 4.4 with $\varepsilon = 0$, and assertion (ii) follows from an application of the same lemma for some $\varepsilon > 0$, with the choice $C_\varepsilon = C^{(\varepsilon/3)} \wedge C(R,M) > 0$. □

5 Proof of Proposition 1.3

In order to show the proposition we have to make sure that the Q_n–typical path does not have too large a range. We know already that the local times are close to the optimal shape, both in the L_1 and L_∞-norms. However, this does not suffice to prove the convergence of the range. Now we have to take care of small pieces of the path that might exceed the optimal range. The main problem is that the contributions to the Q_n-probability of the small and the main part are of different orders. This forces us to develop a machinery that allows to consider these parts separately.

In Section 5.1 we develop the basic tools for cutting paths into two pieces and estimating the Q_n-probability of the whole path in terms of the Q_n-probabilities of the respective parts. It is essential that the local time at the point x where we cut the path is not too big. Otherwise we would make too large an error.

In Section 5.2 we split the event that the path has a large range into three events E_1, E_2 and E_3, which will be dealt with separately. For E_1 the bound in Lemma 4.3(ii) is good enough. For E_2 and E_3 we employ results of Section 5.1 to get improved bounds in Lemmas 5.6 and 5.8.

Finally, in Section 5.3 we finish the proof with a bootstrap method. In a first step we use Lemma 4.3(ii) as well as Lemma 5.6 and 5.8 to show that the range of the path cannot exceed the optimal range $2R^\star\alpha_n$ by more than $R\alpha_n \log\log n$, which is an improvement over the previous $R\alpha_n \log n$ bound of Lemma 4.3(i). In two further steps we apply the intermediate result with a special choice of the parameters in Lemma 4.3(ii), 5.6 and 5.8 again to end up with the final result that the range cannot exceed $2(R^\star + \delta)\alpha_n$ for any $\delta > 0$.

5.1 Surgery on paths

Before we come to the core of the argument in Section 5.2, we develop our main tool: cutting paths into two parts and estimating the parts separately.

The first lemma (Lemma 5.1) is concerned with the basic cutting procedure. We fix a point x at which we would like to cut the path in two pieces. Next we put all loops below x to the first path and all loops above x to a second path. Modulo some bookkeeping for the initial and final part of the original path this is all.

Note that the paths that we obtain from such a crude cutting procedure are rather rough at the boundary point x. That is, around this point we have an exceptionally large decay of the local times and thus an exceptionally large contribution to the Hamiltonian. By adding the mirror images (around x) to the cut paths we get rid of this problem. This yields improved estimates described in a second lemma (Lemma 5.4).

Basic cutting procedure

Now we come to the details. We fix a local time configuration L_n for a path of length n as well as its upstep vector M_n. From this the position $s = s_n$ can be computed. We fix an $x \in \mathbb{N}$ and define the number of loops the path makes above the level x and the number of loops the path makes below the level x:

$$U := M_n(x) - 1\!\!1_{s>x} \qquad \text{and} \qquad D := M_n(x-1) - 1. \tag{5.1}$$

Further let

$$\hat{n} := \sum_{y \geq x} L_n(y) - M_n(x-1) - 1\!\!1_{s>x}. \tag{5.2}$$

The aim is to cut the path into two pieces at the point x. Since there are potentially several possibilities to do so, we have to specify the cutting procedure. We want to cut in such a way that the upstep vectors of the left part is given by $(M_n(y) 1\!\!1_{y \leq x-1})_y$, while the upstep vectors of the part on the right (shifted to the origin) should be $(M_n(y+x) 1\!\!1_{y \geq 0})_y$. The length of the right part is then $\hat{n}$ and the length of the left part is $n - \hat{n}$. The corresponding local times are

$$L'_{n-\hat{n}}(y) = \begin{cases} L_n(y), & \text{if } y < x, \\ M_n(x-1), & \text{if } y = x, \\ 0, & \text{otherwise,} \end{cases} \qquad \text{and} \qquad L''_{\hat{n}}(y) = \begin{cases} L_n(x+y), & \text{if } y > 0, \\ M_n(x) + 1\!\!1_{s \leq x}, & \text{if } y = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

We think of $\hat{n}$ as being small compared with n (although this does not enter the subsequent formulas). Hence we call the path left of x the *major* path and the path right of x the *minor* path.

Lemma 5.1 (Basic cutting estimate) *With the above definitions,*

$$E\big(e^{-H_n} 1\!\!1\{\ell_n = L_n\}\big) = e^{F+G} \times E\big(e^{-H_{n-\hat{n}}} 1\!\!1\{\ell_{n-\hat{n}} = L'_{n-\hat{n}}\}\big) E\big(e^{-H_{\hat{n}}} 1\!\!1\{\ell_{\hat{n}} = L''_{\hat{n}}\}\big), \tag{5.4}$$

where

$$|F| \leq L_n(x)\Big(\beta L_n(x) + \sqrt{\beta H_n(\bar{L}_n)}\Big) \qquad \textit{and} \qquad 0 \leq G \leq L_n(x) \cdot \log 2. \tag{5.5}$$

Proof. Recall the notation in (3.14). Clearly

$$E\big(e^{-H_n} 1\!\!1\{\ell_n = L_n\}\big) = e^{-H_n(\bar{L}_n)} P(\ell_n = L_n).$$

Analogous formulae hold for the expectations on the right hand side of (5.4) Thus we have to estimate the difference of the Hamiltonians and the quotient of the probabilities in (5.4).

We start with the Hamiltonians. Note that the sum in (3.16) differs only in the summands for x and $x+1$. Let $\alpha' = L'_{n-\hat{n}}(x)/L_n(x)$ and $\alpha'' = L''_{\hat{n}}(0)/L_n(x)$. The summand in x is estimated using

$$\begin{aligned}
&\Big|(L_n(x) - L_n(x-1))^2 - \big[L'_{n-\hat{n}}(x)^2 + (L'_{n-\hat{n}}(x-1) - L'_{n-\hat{n}}(x))^2\big]\Big| \\
&\qquad = \Big|\big(1 - 2\alpha'^2\big)L_n(x)^2 - 2(1-\alpha')L_n(x)L_n(x-1)\Big| \\
&\qquad = \big(1 - 2\alpha'(1-\alpha')\big)L_n(x)^2 + 2(1-\alpha')L_n(x)\big|L_n(x-1) - L_n(x)\big| \\
&\qquad \le L_n(x)^2 + 2(1-\alpha')L_n(x)\sqrt{H_n(\bar{L}_n)}\sqrt{\frac{2}{\beta}}.
\end{aligned} \tag{5.6}$$

An analogous formula holds for the summand in $x+1$ with $L''_{\hat{n}}$ instead of $L'_{n-\hat{n}}$. Using the fact that $\alpha' + \alpha'' \ge 1$ we get

$$\big|H_n(\bar{L}_n) - H_{n-\hat{n}}(\bar{L}'_{n-\hat{n}}) - H_{\hat{n}}(\bar{L}''_{\hat{n}})\big| \le L_n(x)\Big(\beta L_n(x) + \sqrt{\beta H_n(\bar{L}_n)}\Big).$$

It remains to estimate the ratio of the probability terms. The original path with local times L_n makes $U+D$ loops from x. There are $\binom{U+D}{U}$ choices for the order of loops above and below x. Hence we get

$$P(\ell_n = L_n) = \binom{U+D}{U} P(\ell_{n-\hat{n}} = L'_{n-\hat{n}})P(\ell_{\hat{n}} = L''_{\hat{n}}).$$

(This formula can be derived also using the matrices P and P^*.) Using $\binom{U+D}{U} \le 2^{U+D} = 2^{L_n(x)-1}$ the proof is completed. □

We continue with two applications of the lemma. The first one estimates the probability for the minor path having a given fixed length. Note that the bound on the Hamiltonian may be inserted freely for large M since the opposite has obviously a negligible probability.

Corollary 5.2 (Estimate for a fixed length of the minor path) *Fix $x \in \mathbb{N}_0$ and $l, k, M \ge 0$ with $l \le k$. Then*

$$\begin{aligned}
&Q_n\Big(\ell_n(x) \le l, \sum_{y \ge x} \ell_n(y) = k, H_n \le \frac{\beta}{2} M\alpha_n \log n\Big) \\
&\qquad\qquad \le n2^l e^{\beta l^2} e^{\sqrt{2\beta M\alpha_n \log n}} \max_{j \in \{0,\dots,l\}} \frac{Z_{n-k+j}Z_{k-j}}{Z_n}.
\end{aligned} \tag{5.7}$$

Proof. Sum (5.4) on all L_n with $L_n(x) \le l$ and $\sum_{y \ge x} L_n(y) = k$ and $H_n(L_n) \le \frac{\beta}{2} M\alpha_n \log n$, and divide by the partition sum Z_n. The extra factor n comes from the fact that here we have at most n possible values for the upstep number $M_n(x)$. □

Recall that we are interested in getting estimates for the Q_n-probability of paths for which the support right of x is large. This was the purpose of Lemma 5.1. For the next corollary of the lemma we introduce the following event: For $x \in \mathbb{N}_0$, $M > 0$ and $k, l, \Lambda \in \mathbb{N}_0$ with $\Lambda \le k$

$$F(x,k,\Lambda,l) = \Big\{\sum_{y \ge x} \ell_n(y) \le k, \ell_n(x+\Lambda) > 0, 0 < \ell_n(x) \le l, H_n \le \frac{\beta}{2} M\alpha_n \log n\Big\}. \tag{5.8}$$

In words, the path visits at least all the integers in $[x, x+\Lambda]$ and does not spend more than k time units right of x. In our later applications, we shall choose k relatively small in comparison to n, such that the main part of the path will indeed be left of x.

Corollary 5.3 (Estimate with a given minimal support of the minor path) *For any x, M, Λ, k, l,*

$$Q_n\big(F(x,k,\Lambda,l)\big) \le kl\cdot 2^l e^{2\beta l^2} e^{2\sqrt{2\beta M\alpha_n\log n}} \max_{\hat n\in\{\Lambda,\dots,k\}} \inf_{\substack{\varphi\in\mathcal{F}_{\hat n}\\ H_n(\varphi)\le\frac{\beta}{2}M\alpha_n\log n}} \frac{Q_{\hat n}(\#\operatorname{supp}\ell_{\hat n}\ge\Lambda)}{Q_{\hat n}(\overline{\ell}_{\hat n}=\varphi)}. \tag{5.9}$$

Proof. From Lemma 5.1 we get the following inequality:

$$Q_n\big(F(x,k,\Lambda,l)\big) \le \frac{1}{Z_n}2^l e^{\beta l^2} e^{\sqrt{2\beta M\alpha_n\log n}}$$
$$\times\sum_{\hat n=\Lambda}^{k}\sum_{b=1}^{l} E\Big(e^{-H_{n-\hat n}}\mathbb{1}\{m_{n-\hat n}(x-1)=b\}\mathbb{1}\{H_{n-\hat n}\le\frac{\beta}{2}M\alpha_n\log n\}\Big)E\Big(e^{-H_{\hat n}}\mathbb{1}\{\#\operatorname{supp}\ell_{\hat n}\ge\Lambda\}\Big). \tag{5.10}$$

Furthermore, Lemma 5.1 implies, for any choice of b, $\hat n$ and for any $\varphi\in\mathcal{F}_{\hat n}$ satisfying $H_{\hat n}(\varphi)\le\frac{\beta}{2}M\alpha_n\log n$,

$$Z_n \ge E\Big(e^{-H_n}\mathbb{1}\{m_n(x-1)=b\}\mathbb{1}\{\ell_n(x)\le l\}\mathbb{1}\{H_n\le\beta M\alpha_n\log n\}\Big)$$
$$\ge E\Big(e^{-H_{n-\hat n}}\mathbb{1}\{m_{n-\hat n}(x-1)=b\}\mathbb{1}\big\{H_{n-\hat n}\le\frac{\beta}{2}M\alpha_n\log n\big\}\Big)E\Big(e^{-\beta H_{\hat n}}\mathbb{1}\{\overline{\ell}_{\hat n}=\varphi\}\Big)e^{-\beta l^2}e^{-\sqrt{2\beta M\alpha_n\log n}}. \tag{5.11}$$

Using (5.11) in (5.10) we get

$$Q_n\big(F(x,k,\Lambda,l)\big) \le 2^l e^{2\beta l^2} e^{2\sqrt{2\beta M\alpha_n\log n}} \sum_{\substack{\Lambda\le\hat n\le k\\ b\le l}} \inf_{\substack{\varphi\in\mathcal{F}_{\hat n}\\ H_{\hat n}(\varphi)\le\frac{\beta}{2}M\alpha_n\log n}} \frac{Q_{\hat n}(\#\operatorname{supp}\ell_{\hat n}\ge\Lambda)}{Q_{\hat n}(\overline{\ell}_{\hat n}=\varphi)}. \tag{5.12}$$

The right hand side is obviously not bigger than the right hand side of (5.9), and this ends the proof. □

Doubling paths

The next lemma is concerned with the following situation: Fix a point $x\in\mathbb{Z}$ and let A be an event that depends only on the local times $\ell_n(y)$ with $y\ge x$ and on which in particular $\sum_{y\ge x}\ell_n(y)=\hat n$ holds for some $\hat n\in\mathbb{N}$ which is much smaller than the length n of the path. We want to estimate $Q_n(A)$ in the situation where the local time $\ell_n(x)$ is too large in order to apply the cutting technique of Lemma 5.1 at the site x. The idea is to add to the piece of the path that lies right of x the mirror image of that path left of x. This enables us to estimate $Q_n(A)$ in terms of the probability that both halves of the concatenated path of length $2\hat n$ are in A.

While the formulation of this lemma requires some notation, the main application, which comes in the subsequent corollary, does not. Thus the reader can skip directly to Corollary 5.5.

Here are the details. For a local time vector L_n denote by $L^x_{\hat n}$ the vector defined by $L^x_{\hat n}(y)=L_n(y+x)\mathbb{1}_{y\ge 0}$. Here $\hat n=\sum_{y\ge x}L_n(y)$. That is, $L^x_{\hat n}$ is the local time vector of the path right of x, but shifted to the origin. Denote by s the endpoint of a path corresponding to L_n. We define $L_{2\hat n}$ by $L_{2\hat n}(((s-x)_+ +\frac12)\pm(y+\frac12))=L_n(x+y)$, $y\ge 0$. This is the local time shape of a path that ends in in the point $2(s-x)_+ +1$. Strictly speaking, we do not require the path of length $2\hat n$ to be symmetric, but rather the resulting local time vector $L_{2\hat n}$ is symmetric.

Assume now that we are given two local time vectors L^1_n and L^2_n and endpoints s^1 and s^2 with $L^1_n(y)=L^2_n(y)=L_n(y)$ for $y\le x$, and $s^1\wedge x=s^2\wedge x=s\wedge x$. The corresponding upstep vectors will always be denoted by M_n^1 respectively M_n^2. Define $s^{1,2} = (s^1 - x)_+ + (s^2 - x)_+ + 1$ and

$$L_{2\hat{n}}^{1,2}(y) = \begin{cases} L_n^{1,x}(x + y - (s^2 - x)_+ - 1), & \text{if } y \geq (s^2 - x)_+ + 1, \\ L_n^{2,x}(x - y + (s^2 - x)_+), & \text{if } y \leq (s^2 - x)_+. \end{cases} \tag{5.13}$$

The pair $(L_n^{1,2}, s^{1,2})$ belongs to paths that end in $s^{1,2}$ and where the second path's mirror image is taken.

Finally let A denote a subset of the local time vectors $L_{\hat{n}}$ for paths of length $\hat{n}$ that stay right of the origin and for which $L_{\hat{n}}(0) = L_n(x)$. Denote

$$A^{1,2} = \left\{L_{2\hat{n}}^{1,2} : \exists x \text{ such that } L_{\hat{n}}^{1,x}, L_{\hat{n}}^{2,x} \in A\right\}.$$

Lemma 5.4 (Doubling paths) *Assume that $B > 0$ is fixed. There exists a constant $C > 0$ depending only on B such that whenever $L_n(x)\alpha_{\hat{n}}/\hat{n} < B$,*

$$Q_n(\ell_{\hat{n}}^x \in A) \leq e^{C\alpha_{\hat{n}} \log \alpha_{\hat{n}}} \sqrt{Q_{2\hat{n}}(\ell_{2\hat{n}} \in A^{1,2})} \tag{5.14}$$

Proof. The basic observation is that for $i = 1, 2$ (recall the notation P_x^s from (3.4)),

$$\frac{Q_n\big(\ell_n = L_n^i,\, S_n = s^i\big)}{Q_n\big(\ell_n = L_n,\, S_n = s\big)} = \prod_{y=x}^{\infty} \Big[\frac{P_y^s\big(M_n^i(y), M_n^i(y-1)\big)}{P_y^s\big(M_n(y), M_n(y-1)\big)} \, \frac{e^{-\beta(L_n^i(y-1) - L_n^i(y))^2}}{e^{-\beta(L_n(y-1) - L_n(y))^2}}\Big]. \tag{5.15}$$

It will be sufficient for our purposes to rewrite the denominators as

$$\begin{aligned} &\prod_{y=x}^{\infty} \big[P_y^s\big(M_n(y), M_n(y-1)\big)\, e^{-\beta(L_n(y-1) - L_n(y))^2}\big] \\ &\qquad = E\big(e^{-H_{2\hat{n}}} 1\!\!1\{\ell_{2\hat{n}} = L_{2\hat{n}}\} 1\!\!1\{S_{2\hat{n}} = 2(s-x)_+ + 1\}\big)^{1/2}. \end{aligned} \tag{5.16}$$

Hence, we have

$$\prod_{i=1,2} \frac{Q_n\big(\ell_n \in A,\, \ell_n(y) = L_n(y),\, y \leq x,\, S_n = s^i\big)}{Q_n\big(\ell_n = L_n,\, S_n = s\big)} \leq \frac{Q_{2\hat{n}}\big(\ell_{2\hat{n}} \in A^{1,2},\, S_{2\hat{n}} = s^{1,2}\big)}{Q_{2\hat{n}}\big(\ell_{2\hat{n}} = L_{2\hat{n}},\, S_{2\hat{n}} = 2(s-x)_+ + 1\big)}. \tag{5.17}$$

We next sum (5.17) over $s^1, s^2 \in \{-n, \dots, n\}$ to arrive at

$$\frac{Q_n\big(\ell_n^x \in A,\, \ell_n(y) = L_n(y),\, y \leq x\big)^2}{Q_n\big(\ell_n = L_n,\, S_n = s\big)^2} \leq n^2 \frac{Q_{2\hat{n}}\big(\ell_{2\hat{n}} \in A^{1,2}\big)}{Q_{2\hat{n}}\big(\ell_{2\hat{n}} = L_{2\hat{n}},\, S_{2\hat{n}} = 2(s-x)_+ + 1\big)}. \tag{5.18}$$

From (5.18), we obtain that

$$Q_n\big(\ell_n^x \in A,\, \ell_n(y) = L_n(y),\, y \leq x\big) \leq Q_n\big(\ell_n = L_n\big) n \sqrt{\frac{Q_{2\hat{n}}\big(\ell_{2\hat{n}} \in A^{1,2}\big)}{\min_{s: L_{2\hat{n}}(s) \geq 1} Q_{2\hat{n}}\big(\ell_{2\hat{n}} = L_{2\hat{n}},\, S_{2\hat{n}} = s\big)}}. \tag{5.19}$$

Summing over all possible values of $L_n(y)$, $y < x$ yields

$$\begin{aligned} Q_n\big(\ell_n^x \in A\big) &\leq Q_n\big(\ell_n^x = L_n^x\big) n \sqrt{\frac{Q_{2\hat{n}}\big(\ell_{2\hat{n}} \in A^{1,2}\big)}{\min_{s: L_{2\hat{n}}(s) \geq 1} Q_{2\hat{n}}\big(\ell_{2\hat{n}} = L_{2\hat{n}},\, S_{2\hat{n}} = s\big)}} \\ &\leq n \sqrt{\frac{Q_{2\hat{n}}\big(\ell_{2\hat{n}} \in A^{1,2}\big)}{\min_{s: L_{2\hat{n}}(s) \geq 1} Q_{2\hat{n}}\big(\ell_{2\hat{n}} = L_{2\hat{n}},\, S_{2\hat{n}} = s\big)}}. \end{aligned} \tag{5.20}$$

Note that in the event in the denominator, it is implicit that $L_{2\hat{n}}$ is symmetric around some point $s + \frac{1}{2}$ and that $L_{2\hat{n}}(s) = L_n(x)$. Now we choose $L_{2\hat{n}}$ such that it is close to a symmetric triangle shape around $s+\frac{1}{2}$ with fixed slope $L_n(x)^2/(2\hat{n})$. Similarly as in the proof of Proposition 1.2 we get that there exists a $C>0$ (depending only on B) such that $Q_{2\hat{n}}(\ell_{2\hat{n}}=L_{2\hat{n}})\geq e^{-2C\alpha_{\hat{n}}\log\hat{n}}$. Indeed, we have that $|L_{2\hat{n}}(x)-L_{2\hat{n}}(x-1)|=L_n(x)^2/(2\hat{n})(1+o(1))$, so that

$$H_{2\hat{n}}=\frac{\beta 2\hat{n}}{L_n(x)}\frac{L_n(x)^4}{4\hat{n}^2}(1+o(1))=\beta\frac{L_n(x)^3}{2\hat{n}^2}(1+o(1)). \tag{5.21}$$

Moreover, uniformly in s such that $L_{2\hat{n}}(s)\geq 1$

$$P(\ell_{2\hat{n}}=L_{2\hat{n}},S_{2\hat{n}}=s)=e^{-\frac{2\hat{n}}{L_n(x)}\log L_n(x)+\mathcal{O}\left(\frac{L_n(x)^2}{\hat{n}}\right)}. \tag{5.22}$$

Hence, using that $L_n(x)\alpha_{\hat{n}}/\hat{n}<B$, we finally obtain that

$$Q_{2\hat{n}}(\ell_{2\hat{n}}=L_{2\hat{n}},\,S_{2\hat{n}}=s)\geq e^{-H_{2\hat{n}}}P(\ell_{2\hat{n}}=L_{2\hat{n}},S_n=s)\geq e^{-C\alpha_{\hat{n}}\log\hat{n}}, \tag{5.23}$$

where $C=C(B)$. Finally, we estimate $n\leq e^{C\alpha_{\hat{n}}\log\hat{n}}$ to arrive at the claim. □

We come to the corollary that was the main reason to state the preceding lemma. The basic splitting procedure gave sufficiently good estimates in a situation where the local times around the cutting point x were small, say of order smaller than $n^{1/8}$. The preceding lemma however gives an estimate that works in the complementary situation where the paths are in

$$A_{x,R}:=\Big\{\#\{y\geq x\colon \ell_n(y)>n^{1/8}\}>\frac{R}{2}\alpha_n\Big\},\qquad x\in\mathbb{Z},\,R>0. \tag{5.24}$$

At this point though, we have to make one more assumption on the paths. They have to be in the set B_x that is defined as follows. For $\eta\in(0,R^\star)$ let $K_\eta=\int_{R^\star-\eta}^{R^\star}\varphi^\star(t)dt$, fix $\varepsilon\in(0,(K_\eta/2)\wedge\varphi^\star(R^\star-\eta))$ and define

$$B_x=\Big\{\Big|\frac{\alpha_n}{n}\ell_n(x)-\varphi^\star(R^\star-\eta)\Big|<\varepsilon\Big\}\cap\Big\{\Big|\frac{1}{n}\sum_{y=x}^{\infty}\ell_n(y)-K_\eta\Big|<\varepsilon\Big\}. \tag{5.25}$$

Corollary 5.5 *There exists $C'>0$ such that for all choices of η, ε and x, for all sufficiently large n and all $R>0$*

$$Q_n(A_{x,R}\cap B_x)\leq n\sup_{n(K_\eta-\varepsilon)\leq\hat{n}\leq n(K_\eta+\varepsilon)}e^{C'\alpha_{\hat{n}}\log\hat{n}}\sqrt{Q_{2\hat{n}}(\#\{\ell_{2\hat{n}}>n^{1/8}\}>2R\alpha_n)}. \tag{5.26}$$

Proof. Note that, for any $n(K_\eta-\varepsilon)\leq\hat{n}\leq n(K_\eta+\varepsilon)$ on the event $A_{x,R}\cap B_x$ for $n>(K_\eta/2)^{-2}$,

$$\frac{\ell_n(x)^2\alpha_{\hat{n}}^2}{\hat{n}^2}\leq\frac{8\varphi^\star(R^\star-\eta)^2}{K_\eta}\left(\frac{\log(nK_\eta/2)}{\log n}\right)^{1/2}\leq\frac{16\varphi^\star(R^\star-\eta)^2}{K_\eta}\leq\frac{3}{R^\star}. \tag{5.27}$$

Hence, we can apply Lemma 5.4. □

5.2 Reducing and splitting the problem

This section sets the stage for the proof of the convergence of the range in Proposition 1.3. First we point out that it is sufficient to consider paths whose local time is close to the optimal one, centered around some site $x^*\in\mathbb{Z}$. Then we split the event under interest into three events E_1, E_2 and E_3. Finally, we give bounds for $Q_n(E_2)$ and $Q_n(E_3)$, respectively. (The term $Q_n(E_1)$ will later be bounded using Lemma 4.3(ii).)

Reducing the problem

Recall that our goal is the derivation of the inequality in (4.9) for all $L > (R^\star+\delta)\alpha_n$ rather than only for $L \geq \text{const}\,\alpha_n \log n$.

Fix $M > 2\chi$ and define the event

$$D_n^\varepsilon(x^*) = \Big\{\|\overline{\ell}_n - \tau_{x^*\alpha_n^{-1}}\varphi^\star\| \leq \varepsilon\Big\} \cap \Big\{H_n \leq M\alpha_n \log n\Big\}, \qquad x^* \in \mathbb{Z}, \varepsilon > 0. \tag{5.28}$$

In words, the local times are close to the optimal shape, centered around the site x^*, and the Hamiltonian is not too large.

Note that, on $D_n^\varepsilon(x^*)$, we have in particular $|\ell_n(x+x^*) - \frac{n}{\alpha_n}\varphi^\star(\frac{x}{\alpha_n})| \leq \varepsilon\frac{n}{\alpha_n}$ for all $x \in \mathbb{Z}$ and $\sum_{x\colon |x-x^*|\leq R^\star\alpha_n} \ell_n(x) \geq (1-\frac{\varepsilon}{2})n$, and hence $|\sum_{x\geq x^*+R^\star\alpha_n} \ell_n(x) - nK_\eta| \leq \varepsilon n$, where $K_\eta = \int_{R^\star-\eta}^{R^\star} \varphi^\star(t)\,dt$. Furthermore, the set $\operatorname{supp}\ell_n = \{S_0,\dots,S_n\}$ contains at least the interval of length $2(R^\star-\varepsilon')\alpha_n$ centered at x^*, where $\varepsilon' = R^\star(1-\sqrt{1-4R^\star\varepsilon/3})$ is chosen such that $\varphi^\star(R^\star-\varepsilon') = \varepsilon$.

Furthermore, by Proposition 1.2, there is $C_\varepsilon > 0$ such that, for sufficiently large $n \in \mathbb{N}$,

$$Q_n\Big(\bigcap_{x^*\in\mathbb{Z}} D_n^\varepsilon(x^*)^{\rm c}\Big) \leq e^{-C_\varepsilon\alpha_n\log n} + e^{-(M-\chi-\varepsilon)\alpha_n\log n} \leq 2e^{-C_\varepsilon\alpha_n\log n}.$$

Because of this and since we need to consider only $x^* \in \{-n,\dots,n\}$ (this is, there are only $2n+1$ choices for x^*, which is small compared with $e^{C_\varepsilon\alpha_n\log n}$), it is sufficient to derive the estimate in (4.9) for $D_n^\varepsilon(x^*) \cap \{\#\operatorname{supp}\ell_n > 2L\}$ instead of $\{\#\operatorname{supp}\ell_n > 2L\}$.

Next we want to make the symmetric problem of bounding the range to a one-sided problem. Note that for any $x^* \in \mathbb{Z}$, on the event $D_n^\varepsilon(x^*)$,

$$\{\#\operatorname{supp}\ell_n > 2L\} \subset \big\{\ell(\lfloor x^*-L\rfloor) > 0\big\} \cup \big\{\ell(\lceil x^*+L\rceil) > 0\big\}.$$

By symmetry, it suffices to consider only one of the two events in the brackets; we shall concentrate on the second one. Hence, to prove Proposition 1.3, it will be sufficient to show, for some $C > 0$, the estimate

$$\max_{x^*\in\mathbb{Z}} Q_n\Big(D_n^\varepsilon(x^*) \cap \big\{\ell(\lceil x^*+L\rceil) > 0\big\}\Big) \leq e^{-C(L-R^\star\alpha_n)}, \qquad L > (R^\star+\delta)\alpha_n \tag{5.29}$$

for all sufficiently large $n \in \mathbb{N}$.

Three events

We split the event in (5.29) in three events E_1, E_2 and E_3 that will be treated separately. Intuitively speaking, E_1 copes the event that we cannot find a good point $x \geq x^* + R^\star\alpha_n$ to make the cutting procedure work, since the local time $\ell_n(x) > n^{1/8}$ is too big. On both E_2 and E_3 there exists at least one potential cutting point $x \geq x^* + R^\star\alpha_n$. On E_2 it has the property that the support right of x is large but the path right of x is short. Finally, on E_3 the path right of x has a considerable length. We will define the events as intersections (E_1) and unions (E_2 and E_3) of the corresponding elementary events $E_1(x)$, $E_2(x)$ and $E_3(x)$, which will be defined next.

For given $R, r > 0$ and $x \in \mathbb{Z}$, define the events

$$A_n(x,R) = \Big\{\ell_n(\lceil x + \tfrac{R}{2}\alpha_n\rceil) > 0\Big\} \qquad \text{and} \qquad B_n(x,r) = \Big\{\sum_{y\geq x} \ell_n(y) \leq nr^2\Big\}. \tag{5.30}$$

In words, on $A_n(x,R)$, the path $(S_0,\dots,S_n)$ also visits the site $\lceil x + R\alpha_n/2\rceil$, and on $B_n(x,r)$, it spends at most nr^2 time units right of x. The parameters r and R will be chosen in the course of the proof.

For $x \geq x^* + R^\star\alpha_n$ define the events

$$\begin{aligned} E_1(x) &= D_n^\varepsilon(x^*) \cap \big\{\ell_n(\lceil x^* + L\rceil) > 0\big\} \cap \Big[\{\ell_n(x) > n^{1/8}\} \cup A_n(x,R)^{\rm c}\Big], \\ E_2(x) &= D_n^\varepsilon(x^*) \cap \Big[\{\ell_n(x) \leq n^{1/8}\} \cap A_n(x,R) \cap B_n(x,r)\Big], \\ E_3(x) &= D_n^\varepsilon(x^*) \cap \Big[\{\ell_n(x) \leq n^{1/8}\} \cap B_n(x,r)^{\rm c}\Big]. \end{aligned} \tag{5.31}$$

Obviously, the event $D_n^\varepsilon(x^*) \cap \{\ell(\lceil x^* + L\rceil) > 0\}$ is contained in $E_1(x) \cup E_2(x) \cup E_3(x)$ for any $x \geq x^* + R^\star\alpha_n$. Hence the de Morgan rules yield that

$$D_n^\varepsilon(x^*) \cap \{\ell(\lceil x^* + L\rceil) > 0\} \subset E_1 \cup E_2 \cup E_3, \tag{5.32}$$

where

$$E_1 = \bigcap_{x \geq x^* + R^\star\alpha_n} E_1(x) \qquad \text{and} \qquad E_i = \bigcup_{x \geq x^* + R^\star\alpha_n} E_i(x) \text{ for } i = 2, 3. \tag{5.33}$$

We will show that for the appropriate values of L, r and R, the probability $Q_n(E_i)$ is small for each $i = 1, 2, 3$, but for different reasons:

- E_1 : If $L \geq (R + R^\star)\alpha_n$, then on $\{\ell_n(\lceil x^* + L\rceil) > 0\}$, the event $A_n(x,R)^{\rm c}$ does not occur for at least $R\alpha_n/2$ values of x right of $R^\star\alpha_n$. Hence, on E_1 at least $R\alpha_n/2$ of the local times are larger than $n^{1/8}$, which is unlikely due to Lemma 4.3(ii).
- E_2 : For some $x \geq x^* + R^\star\alpha_n$, the sum of the local times right of x is less than nr^2, whereas the range is at least $R\alpha_n/2$. If we let $r = r(n)$ decrease with n such that $k = nr^2$ is so small that $R\alpha_n \geq R_0\alpha_k \log k$, then $Q_n(E_2)$ is small due to Lemma 4.3(i).
- E_3 : Since, for some $x \geq x^* + R^\star\alpha_n$, the sum of the local times right of x is at least nr^2, and by cutting the path at x, we can estimate this contribution by the ratio of normalization constants $Z_k Z_{n-k}/Z_n$. Moreover, we know by Proposition 1.2 that $Z_n \approx e^{-\chi\alpha_n \log n}$, and since $\alpha_n \log n$ is quite concave, the above ratio is small.

Bounds for E_2 and E_3

We come to giving bounds for $Q_n(E_2)$ and $Q_n(E_3)$ that will be needed in Section 5.3 to finish the proof of Proposition 1.3.

Lemma 5.6 (Bound for E_2) *For any $C > 0$ and all sufficiently large $n \in \mathbb{N}$, and for all $R, r > 0$, the following implication holds: If $Q_k(\#\operatorname{supp}\ell_k > \frac{R}{2}\alpha_n) \leq e^{-CR\alpha_n}$ holds for all $k \in \mathbb{N}$ with $\frac{R}{2}\alpha_n \leq k \leq nr^2$, then for all $\varepsilon > 0$ and all $x, x^* \in \mathbb{Z}$ with $x \geq x^* + R^\star\alpha_n$,*

$$Q_n(E_2(x)) \leq e^{-\frac{1}{2}CR\alpha_n} e^{R_0\alpha_{nr^2}\log\log(nr^2)}, \tag{5.34}$$

where R_0 is chosen according to Proposition 1.2.

In particular, if $r = r_n$ is small enough such that

$$\alpha_{nr^2} \log\log(nr^2) \leq \frac{CR}{8R_0}\alpha_n, \tag{5.35}$$

then for n large enough

$$Q_n(E_2) \leq e^{-\frac{1}{4}CR\alpha_n}. \tag{5.36}$$

Proof. On the event $E_2(x)$, we cut the local time vector at x into two pieces: the one left and the one right of x. We use Corollary 5.3. This yields

$$Q_n(E_2(x)) \le (nr)^2 2^{n^{1/8}} e^{2\beta n^{1/4}} e^{2\beta\sqrt{M}\,(n\log n)^{3/8}} \max_{\frac{R}{2}\alpha_n \le k \le nr^2} \inf_{\substack{\varphi\in\mathcal{F}_k \\ H_k(\varphi)\le M\alpha_n \log n}} \frac{Q_k(\#\operatorname{supp}\ell_k > \frac{R}{2}\alpha_n)}{Q_k(\overline{\ell}_k = \varphi)} \tag{5.37}$$

Note that from Lemma 3.4 and Proposition 1.2(i) we get that for some R_0 and for sufficiently large n, and for all $\frac{R}{2}\alpha_n \le k \le nr^2$,

$$\sup_{\substack{\varphi\in\mathcal{F}_k \\ H_k(\varphi)\le M\alpha_n \log n}} Q_k(\overline{\ell}_k = \varphi) \ge e^{-R_0\alpha_k \log\log k}. \tag{5.38}$$

We substitute (5.38) in (5.37), use the assumption and the fact that for large n

$$(nr)^2 2^{n^{1/8}} e^{2\beta n^{1/4}} e^{2\beta\sqrt{M}\,(n\log n)^{3/8}} \le e^{\frac{1}{2}CR\alpha_n},$$

to get

$$Q_n(E_2(x)) \le e^{-\frac{1}{2}CR\alpha_n} \max_{\frac{R}{2}\alpha_n \le k \le nr^2} e^{R_0\alpha_k \log\log k}.$$

Now use the monotonicity of $k \mapsto \alpha_k \log\log k$ to arrive at the estimate (5.34).

If (5.35) holds, then (5.34) clearly implies, for sufficiently large n and all $x, x^* \in \mathbb{Z}$ with $x \ge x^* + R^\star\alpha_n$,

$$Q_n(E_2(x)) \le e^{-\frac{3}{8}CR\alpha_n}.$$

Now use that $Q_n(E_2(x)) \ne 0$ for at most $2n+1$ values of x to get

$$Q_n(E_2) \le \sum_{x\ge x^*+R^\star\alpha_n} Q_n(E_2(x)) \le (2n+1)e^{-\frac{3}{8}CR\alpha_n} \le e^{-\frac{1}{4}CR\alpha_n}$$

for n large enough. □

Abbreviate $f(x) = \sqrt{x}\log^{\frac{3}{4}} x$ for any $n \in \mathbb{N}$, so that $f(n) = \alpha_n \log n$.

Lemma 5.7 (A concavity bound) *If $n \ge 10000$, and $k \le n/2$*

$$\frac{1}{2}f(k) + f(n-k) \ge f(n).$$

Proof. Let $g_n(x) = \frac{1}{2}f(x) + f(n-x) - f(n)$. Since $f''(x) = -\frac{1}{16}x^{-3/2}(\log x)^{-5/4}[4\log^2 x + 3] < 0$ and since $g_n''(x) = \frac{1}{2}f''(x) + f''(n-x)$, we see that g_n is concave on $[1, n-1]$. Furthermore, one can compute that $g_n(2) > 0$ for $n \ge 9$, and that $g_n(n/2) = \frac{3}{2}f(n/2) - f(n)$ which is is positive for $n \ge 9687$. Hence, g_n must be strictly positive between 2 and $\frac{n}{2}$. □

Lemma 5.8 (Bound for E_3) *Fix $\varepsilon \in (0, \frac{1}{2})$. Then for any sufficiently large $n \in \mathbb{N}$, for any $r > 0$ such that $n^{1/8} \le nr^2/2$, and for any $x, x^* \in \mathbb{Z}$ with $x \ge x^* + R^\star\alpha_n$,*

$$Q_n(E_3(x)) \le e^{-\frac{1}{2}\chi f(nr^2/2)} e^{5R_0\alpha_n \log\log n}, \tag{5.39}$$

where R_0 is chosen according to Proposition 1.2.

Proof. On the event $E_3(x)$ we cut the local time vector at x into two pieces: the one left and the one right of x. We use Corollary 5.2. This yields

$$Q_n(E_3(x)) \le n^2 \max_{j\le n^{1/8}} \max_{nr^2\le k\le n\varepsilon} \frac{Z_{k-j}\, Z_{n-k+j}}{Z_n} e^{R_0\alpha_n \log\log n}.$$

Now we use Proposition 1.2(i) for $k-j$, $n-k+j$ and n and recall the abbreviation $f(n) = \alpha_n \log n$. Hence, we obtain

$$Q_n(E_3(x)) \le n^2 e^{4R_0\alpha_n \log\log n} \max_{j\le n^{1/8}} \max_{nr^2\le k\le n\varepsilon} e^{-\chi[f(k-j)+f(n-k+j)-f(n)]}.$$

Now use Lemma 5.7, the monotonicity of $n\mapsto f(n)$ together with our assumption $n^{1/8} \le nr^2/2$ to obtain that

$$f(k-j)+f(n-k+j)-f(n) \ge \frac{1}{2}f(k-j) \ge \frac{1}{2}f(nr^2-n^{1/8}) \ge \frac{1}{2}f(nr^2/2).$$

Furthermore, estimate $n^2 \le e^{R_0\alpha_n \log\log n}$ to arrive at (5.39). □

5.3 Proof of Proposition 1.3: the bootstrap

Recall that the proof of Proposition 1.3 follows from an extension of the inequality in (4.9) from all $L > R_0\alpha_n \log n$ to all $L > (R^\star+\delta)\alpha_n$ for any $\delta > 0$. In the view of Section 5.2, it is sufficient to give the respective bounds for $Q_n(E_1)$, $Q_n(E_2)$ and $Q_n(E_3)$ for this choice of L.

The bound on $Q_n(E_2)$ is the most difficult one. We use a two-step bootstrap method for proving it. More precisely, in Step 1 below we use (4.9) for $L \ge \mathrm{const}\,\alpha_n \log n$ with an appropriate choice of the parameters in order to extend its validity to all L larger than $\mathrm{const}\,\alpha_n \log\log n$. In the following Steps 2 and 3, we in turn use this improved bound (with another appropriate choice of the parameters) in order to prove (4.9) for all $L \ge (R^\star+\delta)\alpha_n$. The latter improvement finally implies Proposition 1.3.

STEP 1 *There exist $C, R_1 > 0$ such that for all sufficiently large $n\in\mathbb{N}$ and all $R > R_1 \log\log n$*

$$Q_n\big(\#\operatorname{supp}\ell_n > 2(R+R^\star)\alpha_n\big) \le e^{-CR\alpha_n}. \tag{5.40}$$

Proof. We pick the parameters as $L = (R+R^\star)\alpha_n$ and $r = \frac{R}{4R_0\log n}$ where R_0 is chosen so large that Proposition 1.2 and Lemma 4.3 hold with this R_0, and R is as in the claim. Recall from (5.29) and (5.32) that it is sufficient to prove that $Q_n(E_1)$, $Q_n(E_2)$ and $Q_n(E_3)$ each do not exceed $e^{-CR\alpha_n}$ for sufficiently large n, for some $\varepsilon > 0$ and some $C > 0$.

Pick $\varepsilon > 0$ small enough in order to apply Lemma 5.8. In the following, we estimate the probabilities of E_1, E_2 and E_3 separately.

$\mathbf{E_1}$: Note that on the set E_1, we have $\ell_n(x) > n^{1/8}$ for at least $\frac{R}{2}\alpha_n$ sites $x\in\mathbb{Z}$. Indeed, on $\{\ell_n(\lceil x^* + L\rceil) > 0\}$, for at least $\frac{R}{2}\alpha_n$ sites x, the event $A_n(x,R)^{\mathrm{c}}$ does not occur, hence the event $\{\ell_n(x) > n^{1/8}\}$ must occur for these sites x.

Therefore, according to Lemma 4.3(ii), we have, for sufficiently large $n\in\mathbb{N}$,

$$Q_n(E_1) \le Q_n\Big(\#\{\ell_n > n^{1/8}\} > \frac{R}{2}\alpha_n\Big) \le e^{-\frac{1}{2}CR\alpha_n \log n}.$$

$\mathbf{E_2}$: We apply Lemma 5.6. First we check that the assumption of that lemma is satisfied. For doing this, we apply Lemma 4.3. Note that, for sufficiently large n, we have $R\alpha_n \ge \frac{R}{2}\alpha_n(1+o(1)) = 2R_0\alpha_{nr^2}\log(nr^2) \ge 2R_0\alpha_k \log k$ if $k \le nr^2$. Therefore, (4.9) implies that for any k with $\frac{R}{2}\alpha_n \le k \le nr^2$

$$Q_k\big(\#\operatorname{supp}\ell_k > \tfrac{R}{2}\alpha_n\big) \le e^{-CR\alpha_n/4}.$$

Note that with our choices of r and R,

$$\alpha_k \le \alpha_n \frac{R}{4R_0 \log n} \qquad \text{if } k \le nr^2.$$

Hence, $\alpha_{nr^2} \log\log(nr^2) = o(R\alpha_n)$. Therefore, Lemma 5.6 yields that $Q_n(E_2) \le e^{-\frac{1}{16}CR\alpha_n}$ for sufficiently large n.

$\mathbf{E_3}$: We apply Lemma 5.8 and obtain, for any $x^* \in \mathbb{Z}$,

$$Q_n(E_3) \le \sum_{x \ge x^* + R^\star \alpha_n} Q_n(E_3(x)) \le (2n+1) e^{-\frac{1}{2}\chi \alpha_{nr^2/2} \log(nr^2/2)} e^{5R_0\alpha_n \log\log n}.$$

Now use that $\alpha_{nr^2/2} \log(nr^2/2) = \frac{R}{4\sqrt{2}R_0} \alpha_n (1+o(1))$ and that $\log\log n \le \frac{R}{R_1}$, according to our choices of r and R. If we choose R_1 large enough, then we obtain therefore the estimate $Q_n(E_3) \le e^{-\frac{\chi}{16R_0} R\alpha_n}$ for any sufficiently large n.

Collecting the estimates for E_1, E_2 and E_3 we obtain the assertion with an appropriate choice of C. □

At this point, we have shown that the event that the local time support exceeds $R_0\alpha_n \log\log n$ is negligible under Q_n. We will next improve this result to the result that the size of the support converges to $2R^\star\alpha_n$. For this, we will use the result of Step 1.

For the following two steps, we put $L = (R + R^\star)\alpha_n$. Recall (5.32).

STEP 2 *There is $C > 0$ such that, for any $\delta > 0$ and any sufficiently small $\varepsilon > 0$ and sufficiently large $n \in \mathbb{N}$, for any $R \ge \delta$ and any $x^* \in \mathbb{Z}$,*

$$Q_n(E_1) \le e^{-CR\alpha_n \log n}. \tag{5.41}$$

Proof. Pick $C' > 0$ according to Corollary 5.5 and $C > 0$ according to Lemma 4.3. Let $\delta > 0$ be given and choose $\eta \in (0, \delta)$ so small that $f(n(K_\eta + \varepsilon)) \le \frac{CR}{2C'} f(n)$ for all $\varepsilon \in (0, K_\eta)$ and all sufficiently large n, where we recall that $f(n) = \alpha_n \log n$ and $K_\eta = \int_{R^\star - \eta}^{R^\star} \varphi^\star(t)\, dt$. Let ε be in $(0, \frac{K_\eta}{2} \wedge \varphi^\star(R^\star - \eta))$.

We split the local times vector at the site $x = \lfloor x^* + (R^\star - \eta)\alpha_n \rfloor$. On the event E_1 we have $\ell_n(y) > n^{1/8}$ at least for every $y \in \mathbb{Z}$ with $0 \le y - x^* - R^\star\alpha_n \le \frac{R}{2}\alpha_n$. Furthermore, recall from the beginning of Section 5.2 that, on E_1, the number $l = \ell_n(x)$ lies inbetween $\frac{n}{\alpha_n}[\varphi^\star(R^\star - \eta) - \varepsilon]$ and $\frac{n}{\alpha_n}[\varphi^\star(R^\star - \eta) + \varepsilon]$ and that $k = \sum_{y \ge x} \ell_n(y)$ lies inbetween $n(K_\eta - \varepsilon)$ and $n(K_\eta + \varepsilon)$. Hence, $E_1 \subset A_{x,R} \cap B_x$ for the events defined in (5.24) and (5.25). Use Corollary 5.5 and Lemma 4.3(ii) to obtain the estimate

$$\begin{aligned} Q_n(E_1) &\le \max_{n(K_\eta - \varepsilon) \le k \le n(K_\eta + \varepsilon)} e^{C'\alpha_k \log k} \sqrt{Q_{2k}\big(\#\{\ell_{2k} > n^{1/8}\} > R\alpha_n\big)} \\ &\le e^{C' f(n(K_\eta + \varepsilon))} \max_{n(K_\eta - \varepsilon) \le k \le n(K_\eta + \varepsilon)} \sqrt{Q_{2k}\big(\#\{\ell_{2k} > (2k)^{1/8}\} > \tfrac{R\alpha_n}{\alpha_{2k}} \alpha_{2k}\big)} \\ &\le e^{CRf(n)/2} \max_{n(K_\eta - \varepsilon) \le k \le n(K_\eta + \varepsilon)} e^{-CR\alpha_n \log(2k)} \\ &\le e^{\frac{CR}{2}\alpha_n \log n} e^{-CR\alpha_n \log(2n(K_\eta - \varepsilon))}, \end{aligned} \tag{5.42}$$

and this upper bound does not exceed $e^{-\frac{C}{4}R\alpha_n \log n}$ for sufficiently large n, depending on η and ε. □

Finally, from the assertion of the following last step, Proposition 1.3 follows.

STEP 3 *There is $C>0$ such that, for any $\delta>0$, any sufficiently large $n\in\mathbb{N}$ and any $R\geq\delta$,*

$$Q_n\big(\#\operatorname{supp}\ell_n > 2(R+R^\star)\alpha_n\big)\leq e^{-CR\alpha_n}. \tag{5.43}$$

Proof. This time we choose $L=(R+R^\star)\alpha_n$ and

$$r=\frac{R}{16(R_1+(R_0/C))\log\log n},$$

where C and R_1 are the constants from Step 1, and R_0 is the constant from Proposition 1.2. Again recall from (5.29) and (5.32) that it is enough to prove that $Q_n(E_1)$, $Q_n(E_2)$ and $Q_n(E_3)$ each do not exceed $e^{-CR\alpha_n}$ for some $C>0$ and all sufficiently large n.

Pick $\varepsilon>0$ so small that Lemma 5.8 and Steps 1 and 2 apply. In Step 2 we have already estimated $Q_n(E_1)\leq e^{-CR\alpha_n\log n}$ for large n and some $C>0$. We only have to handle $Q_n(E_2)$ and $Q_n(E_3)$.

E$_2$: We want to apply Lemma 5.6 with the above choices of R and r. In order to check the assumption of Lemma 5.6, we use Step 1 for nr^2 instead of n. For n large enough $R\alpha_n\geq 8R^\star\alpha_{nr^2}$. Thus for n large enough and $k\leq nr^2$

$$\frac{R\alpha_n}{4\alpha_k}-R^\star\geq\frac{R}{8}\frac{\alpha_n}{\alpha_k}\geq\frac{R}{8}\frac{\alpha_n}{\alpha_{nr^2}}\geq\frac{R}{16r}\geq R_1\log\log n.$$

Hence we can apply Step 1 (with $\frac{R}{4}\frac{\alpha_n}{\alpha_k}-R^\star$ instead of R) to get that there exist K_0 and N_0 such that for $k\geq K_0$ and $n\geq N_0$ with $k\leq nr^2$

$$Q_k\left(\#\operatorname{supp}\ell_k>\frac{R}{2}\alpha_n\right)=Q_k\left(\#\operatorname{supp}\ell_k>2\Big(\Big(\frac{R}{4}\frac{\alpha_n}{\alpha_k}-R^\star\Big)+R^\star\Big)\alpha_k\right)$$
$$\leq e^{-C(\frac{R}{4}\frac{\alpha_n}{\alpha_k}-R^\star)\alpha_k}\leq e^{-\frac{C}{4}R\alpha_n+CR^\star\alpha_{nr^2}}\leq e^{-\frac{1}{8}CR\alpha_n}.$$

Hence, we may apply Lemma 5.6. Note that by assumption on r, for sufficiently large n,

$$\alpha_{nr^2}\log\log(nr^2)\leq 2r\alpha_n\log\log n\leq\frac{CR\alpha_n}{8R_0}.$$

Thus Lemma 5.6 yields that for $n\geq N_0$ large enough such that $\frac{R}{2}\alpha_n\geq K_0$ we have

$$Q_n(E_2)\leq e^{-\frac{1}{16}CR\alpha_n}.$$

E$_3$: We apply Lemma 5.8 to obtain

$$Q_n(E_3)\leq\sum_{x\geq x^*+R^\star\alpha_n}Q_n(E_3(x))\leq(2n+1)e^{-\frac{1}{2}\chi f(nr^2/2)}\,e^{5R_0\alpha_n\log\log n}.$$

Now use that

$$f(nr^2/2)=\alpha_{nr^2/2}\log(nr^2/2)=\frac{R}{16\sqrt{2}(R_1+(R_0/C))}\alpha_n\frac{\log n}{\log\log n}(1+o(1))$$

to see that we have even the bound $Q_n(E_3)\leq e^{-CR\alpha_n\frac{\log n}{\log\log n}}$ for all sufficiently large n and some $C>0$ sufficiently small.

Collecting all the estimates for $Q_n(E_1)$, $Q_n(E_2)$ and $Q_n(E_3)$, we arrive at the assertion with an appropriate choice of C. □

Acknowledgment. The work of R.v.d.H. and W.K. was carried out in part at the Fields Institute, Toronto. The work of W.K. was partially supported by "Don Dawson's Max Planck Award for International Cooperation". We are grateful to EURANDOM in Eindhoven, the Netherlands, for offering us the opportunity to work there.

References

[Fl71] P.J. FLORY, *Principles of Polymer Chemistry*, Cornell University Press, Ithaca NY, 1971.

[HHK97] R. VAN DER HOFSTAD, F. DEN HOLLANDER and W. KÖNIG, *Central limit theorem for a weakly interacting random polymer*, Markov Processes Relat. Fields 3, 1-62, 1997.

[HKl01] R. VAN DER HOFSTAD and A. KLENKE, *Self-attractive random polymers*, *to appear in Ann. Appl. Probab., 2001.*

[HKö01] R. VAN DER HOFSTAD and W. KÖNIG, *A survey of one-dimensional random polymers, to appear in Jour. Stat. Phys., 2001.*

[dH96] F. DEN HOLLANDER, *Random polymers,* Stat. Nederl. 50, 136-145, 1996.

[Kn63] F.B. KNIGHT, *Random walks and a sojourn density process of Brownian motion,* Transactions of the AMS 109, 56-86, 1963.

[Sz98] A.-S. SZNITMAN, *Brownian Motion, Obstacles and Random Media.* Springer, Berlin, 1998.

[Va98] C. VANDERZANDE, *Lattice Models of Polymers*, Cambridge University Press, Cambridge, 1998.