Scaling limits for the random walk penalized by its range in dimension one

Nicolas Bouchot

Scaling limits for the random walk penalized by its range in dimension one

Nicolas Bouchot

February 23, 2022

Abstract

In this article we study a one dimensional model for a polymer in a poor solvent: the random walk on $\mathbb{Z}$ penalized by its range. More precisely, we consider a Gibbs transformation of the law of the simple symmmetric random walk by a weight $\exp(-h_n|\mathcal{R}_n|)$, with $|\mathcal{R}_n|$ the number of visited sites and h_n a size-dependent positive parameter. We use gambler's ruin estimates to obtain exact asymptotics for the partition function, that enables us to obtain a precise description of trajectories, in particular scaling limits for the center and the amplitude of the range. A phase transition for the fluctuations around an optimal amplitude is identified at $h_n \asymp n^{1/4}$, inherent to the underlying lattice structure.

1 Introduction of the model and main results

Consider a simple symmetric random walk $(S_k)_{k\geq 0}$ on $\mathbb{Z}^d$, $d \geq 1$, starting from 0, with law denoted $\mathbb{P}$. For $h > 0$, we define the following Gibbs transformation of $\mathbb{P}$, called the *polymer measure*

$$\mathrm{d}\mathbb{P}_{n,h}(S) = \frac{1}{Z_{n,h}}e^{-h|\mathcal{R}_n(S)|}\mathrm{d}\mathbb{P}(S),$$

where $\mathcal{R}_n(S) := \{S_0, \ldots, S_n\}$ is the range of the random walk up to time n and $|\cdot|$ is the cardinal measure. The normalizing quantity

$$Z_{n,h} = \mathbb{E}\left[e^{-h|\mathcal{R}_n(S)|}\right]$$

is called the *partition function* and is such that $\mathbb{P}_{n,h}$ is a probability measure on the space of trajectories of length n.

In any dimension $d \geq 1$, the asymptotics for the log-partition function are known since Donsker and Varadhan [12]. These asymptotics strongly suggest that a polymer of length will typically fold (and fill) into a ball of radius $\rho\, n^{\frac{1}{d+2}}$ for some specific constant $\rho = \rho(d, h)$. This has been proved by [5] in dimension $d = 2$, but only much more recently in dimension $d \geq 3$, by Berestycki and Cerf [1] and Ding, Fukushima, Sun and Xu [10]. More precisely, for $h = 1$ (easily generalized to any $h > 0$), they prove that there exists a positive ρ_d, which only depends on the dimension d, such that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_{n,1}\left(\exists x \in \mathbb{R}^d, B\big(x, (1-\varepsilon)\rho_d n^{\frac{1}{d+2}}\big) \cap \mathbb{Z}^d \subset \mathcal{R}_n \subset B\big(x, (1+\varepsilon)\rho_d n^{\frac{1}{d+2}}\big)\right) = 1\,,$$

1

where $B(x, r)$ is the d-dimensional Euclidean ball centered at x of radius r.

In dimension $d = 1$, this is much easier since the range is uniquely determined by its two endpoints (and always fills completely the one-dimensional ball). This allows for more explicit calculations using mostly gambler's ruin estimates. In particular, one easily derives that $n^{-1/3}|\mathcal{R}_n|$ converges to $\left(\frac{\pi^2}{h}\right)^{1/3}$ in $\mathbb{P}_{n,h}$-probability.

1.1 Outline of the paper

In the rest of the work, we focus only on the case of dimension $d = 1$. Also, we allow the penalization intensity to depend on the length of the polymer, meaning $h = h_n$ now depends on n. We exploit gambler's ruin estimates to their full potential and derive exact asymptotics for the partition function (not only for the log-partition function). Afterwards, we will be able to prove a scaling limit (actually we prove a local limit theorem) for the joint law of the center W_n and the amplitude T_n of the range:

$$T_n := \max_{k \leq n} S_k - \min_{k \leq n} S_k = |\mathcal{R}_n| - 1, \qquad W_n := \frac{T_n}{2} + \min_{k \leq n} S_k = \frac{1}{2}\left(\max_{k \leq n} S_k + \min_{k \leq n} S_k\right).$$

For the sake of the exposition, let us consider the case

$$\lim_{n \to \infty} n^{-\gamma} h_n = \hat{h} \in (0, +\infty), \qquad \text{for some } \gamma \in \mathbb{R}. \tag{1.1}$$

Some results are already present in [3] (which considers a disordered version of the model):

(i) if $\gamma < -\frac{1}{2}$ then $\mathbb{P}_{n,h_n}$ is close in total variation to $\mathbb{P}$;

(ii) if $\gamma \in (-\frac{1}{2}, 1)$ then $\left(\frac{n\pi^2}{h_n}\right)^{1/3} T_n$ converges to 1 in $\mathbb{P}_{n,h_n}$-probability.

(iii) if $\gamma > 1$ then $\mathbb{P}_{n,h_n}$ is concentrated on trajectories visiting only two sites.

Since cases (i) and (iii) are degenerate, we focus on the case $\gamma \in (-\frac{1}{2}, 1)$. In this paper, we give another proof of the convergence $\left(\frac{n\pi^2}{h_n}\right)^{1/3} T_n \to 1$ and we additionally identify the fluctuations of $T_n - \left(\frac{n\pi^2}{h_n}\right)^{1/3}$. We find that a phase transition occurs at $\gamma = \frac{1}{4}$ for the fluctuations:

(i) if $\gamma < \frac{1}{4}$ the fluctuations, normalized by $\left(\frac{n}{h_n^4}\right)^{1/6}$, converge to a Gaussian variable;

(ii) if $\gamma > \frac{1}{4}$ the range penalization is strong enough to collapse the range on $\left(\frac{n\pi^2}{h_n}\right)^{1/3}$ in the sense that the fluctuations live on a finite set (of cardinality 1, sometimes 2).

We will also prove that $\left(\frac{n\pi^2}{h_n}\right)^{-1/3} W_n$ converges to a variable with density $\frac{\pi}{2}\cos(u)\mathbb{1}_{[-\frac{1}{2}, \frac{1}{2}]}(u)$ with respect to the Lebesgue measure and is independent from the fluctuations. This type of results appears to be folklore for confined polymers, as the density is the eigenfunction associated with the principal Dirichlet eigenvalue of the Laplacian on $[0, 1]$ (see e.g. [9, Ch. 8]), but we are not aware of a proof written in detail (at least for the random walk penalized by its range).

Notation. In the rest of the paper we shall use the standard notation: as $x \to a$, we write $g(x) \sim f(x)$ if $\lim_{x \to a} \frac{g(x)}{f(x)} = 1$, $g(x) = \bar{o}(f(x))$ if $\lim_{x \to a} \frac{g(x)}{f(x)} = 0$, $g(x) = \bar{\mathcal{O}}(f(x))$ if $\limsup_{x \to a}\left|\frac{g(x)}{f(x)}\right| < +\infty$ and $f \asymp g$ if $g(x) = \bar{\mathcal{O}}(f(x))$ and $f(x) = \bar{\mathcal{O}}(g(x))$.

We also extensively use the following notation: for $\mathcal{A}$ an event, we denote

$$Z_{n,h_n}(\mathcal{A}) := \mathbb{E}\left[e^{-h_n|\mathcal{R}_n(S)|}\mathbb{1}_{\{S \in \mathcal{A}\}}\right],$$

so that in particular $\mathbb{P}_{n,h_n}(\mathcal{A}) = \frac{1}{Z_{n,h_n}} Z_{n,h_n}(\mathcal{A})$.

1.2 Main results

The following two theorems summarize our results, the first being the main result regarding the asymptotic behavior of (T_n, W_n) and the second being asymptotics for Z_{n,h_n} that have a use of their own.

We define the following quantities, that will be used throughout the paper:

$$T_n^* = T_n^*(h_n) := \left(\frac{n\pi^2}{h_n}\right)^{1/3}, \qquad a_n = a_n(h_n) := \frac{1}{\sqrt{3}}\left(\frac{n\pi^2}{h_n^4}\right)^{1/6} = \frac{1}{\sqrt{3n\pi^2}}(T_n^*)^2. \quad (1.2)$$

Note that $\lim_{n\to\infty} a_n = +\infty$ if and only if $\lim_{n\to\infty} n^{-1/4}h_n = 0$.

Theorem 1.1. • *Assume that $h_n \geq n^{-1/2}(\log n)^{3/2}$ and $\lim_{n\to\infty} n^{-1/4}h_n = 0$; in other words $\gamma \in (-\frac{1}{2}, \frac{1}{4})$ in (1.1). Then under $\mathbb{P}_{n,h_n}$, we have the following convergence in distribution*

$$\left(\frac{T_n - T_n^*}{a_n} ; \frac{W_n}{T_n^*}\right) \xrightarrow[n\to+\infty]{(d)} (\mathcal{T}, \mathcal{W}),$$

where the random variables $\mathcal{T}$ and $\mathcal{W}$ are independent with $\mathcal{T} \sim \mathcal{N}(0,1)$ and $\mathcal{W}$ with density given by $\frac{\pi}{2}\cos(\pi u)\mathbb{1}_{[-\frac{1}{2}, \frac{1}{2}]}(u)$.

• *Assume that $\lim_{n\to\infty} n^{-1/4}h_n = +\infty$ and $\lim_{n\to\infty} n^{-1}h_n = 0$; in other words, $\gamma \in (\frac{1}{4}, 1)$ in (1.1). Denote t_n^o the decimal part of $T_n^* - 2$ and define $\mathcal{A}_n$ as $\{0\}$ if $t_n^o < \frac{1}{2}$, $\{1\}$ if $t_n^o > \frac{1}{2}$ and $\{0,1\}$ if $t_n^o = \frac{1}{2}$. Then we have*

$$\lim_{n\to\infty} \mathbb{P}_{n,h_n}\left(T_n - \lfloor T_n^* - 2\rfloor \notin \mathcal{A}_n\right) = 0$$

Also, under $\mathbb{P}_{n,h_n}$ we have the convergence in distribution $\frac{W_n}{T_n^*} \xrightarrow{(d)} \mathcal{W}$.

Comment. The term $a_n = \frac{1}{\sqrt{3n\pi^2}}(T_n^*)^2$ in Theorem 1.1 arises naturally as a Taylor expansion coefficient in the exponential part of the partition function after injecting gambler's ruin formulae, see Section 1.4 below.

Theorem 1.2. *We have the following exact asymptotics:*
• *Assume that $h_n \geq n^{-1/2}(\log n)^{3/2}$ and $\lim_{n\to\infty} n^{-1/4}h_n = 0$; in other words, $\gamma \in (-\frac{1}{2}, \frac{1}{4})$ in (1.1). Then, as $n \to \infty$*

$$Z_{n,h_n} = (1 + \bar{o}(1))\frac{16\sqrt{2}}{\sqrt{3\pi}}\left(\frac{\cosh(h_n) - 1}{h_n}\right)\sqrt{n} \exp\left(-\frac{3}{2}h_n T_n^*\right).$$

• *Assume that $\lim_{n\to\infty} n^{-1/4}h_n = +\infty$ and $\lim_{n\to\infty} n^{-1}h_n = 0$; in other words, $\gamma \in (\frac{1}{4}, 1)$ in (1.1). Denote $\delta_n = \mathbb{1}_{\{t_n^o \geq \frac{1}{2}\}} - t_n^o$ with t_n^o the decimal part of $T_n^* - 2$. Then, as $n \to \infty$*

$$Z_{n,h_n} = \frac{16}{\pi^{4/3}}(1 + \mathbb{1}_{\{t_n^o = \frac{1}{2}\}})\left(\frac{\cosh(h_n) - 1}{h_n^{1/3}}\right)n^{1/3} \exp\left(-\frac{3}{2}h_n T_n^* - \left(6 + \frac{\pi^2}{12} + \frac{3}{2}\delta_n^2\right)\frac{\pi n}{(T_n^*)^4}(1 + \bar{o}(1))\right).$$

Note that $\lim_{n\to\infty} \frac{n}{(T_n^*)^4} = +\infty$ so that $\bar{o}(\frac{n}{(T_n^*)^4})$ could still diverge.

For the sake of completeness, we add the following result concerning the critical case $\lim_{n\to\infty} n^{-1/4}h_n = \hat{h} \in (0, +\infty)$.

3

Proposition 1.3. *Suppose that* $\lim\limits_{n\to\infty} n^{-1/4}h_n = \hat{h} \in (0,+\infty)$, *so in particular we have* $\lim_{n\to\infty} a_n = \frac{\pi^{5/6}}{\sqrt{3}\hat{h}^{2/3}} =: a$. *Then, as* $n \to \infty$, *we have*

$$Z_{n,h_n} = (1 + \bar{o}(1))\frac{16}{\pi^{4/3}}\left(\frac{\cosh(h_n) - 1}{h_n^{1/3}}\right)n^{1/3}e^{-\frac{3}{2}h_n T_n^*}\theta_n(a), \quad \text{with } \theta_n(a) := \sum_{t=-\infty}^{+\infty} e^{-\frac{(t-\delta_n)^2}{a^2}}$$

(recall that $\delta_n = \mathbb{1}_{\{t_n^o \geq \frac{1}{2}\}} - t_n^o$ *with* t_n^o *the decimal part of* $T_n^* - 2$*). Furthermore, for any integers* $r \leq s$, *as* $n \to \infty$ *we have*

$$\mathbb{P}_{n,h_n}\left(r \leq T_n - \lfloor T_n^* - 2\rfloor \leq s\right) = (1 + \bar{o}(1))\frac{1}{\theta_n(a)}\sum_{t=r}^{s} e^{-\frac{(t-\delta_n)^2}{a^2}}.$$

1.3 Range's endpoints and confinement estimates

Let us now state some estimates for the probability that the range of a random walk is exactly a given interval. The proof is postponed to Section 4 and follows from gambler's ruin estimates that can be found in [13, Chap. XIV].

Let x, y be two non-negative integers and denote by $E_x^y(n)$ the following event

$$E_x^y(n) := \{\mathcal{R}_n = [\![-x, y]\!]\} = \{M_n^- = -x, \, M_n^+ = y\},$$

where we also introduced $M_n^- := \min_{k \leq n} S_k$, $M_n^+ := \max_{k \leq n} S_k$, and used the standard notation $[\![a, b]\!] = [a, b] \cap \mathbb{Z}$. We also define the following function g, that encodes the exponential decay rate of confinement probabilities inside a strip:

$$g(T) := -\log\cos\left(\frac{\pi}{T}\right) = \frac{\pi^2}{2T^2} + \frac{\pi^4}{12T^4} + \bar{\mathcal{O}}(T^{-6}) \quad \text{as } T \to \infty. \tag{1.3}$$

The main result used in the rest of the paper is the following. It is based on sharp gambler's ruin estimates, see Lemmas 4.2-4.3 below.

Theorem 1.4. *Let* x, y *be two positive integers and denote* $T := x + y$. *Then, as* $n \to \infty$:

- *If* $\lim_{n\to\infty} \frac{n}{T^3} = +\infty$, *then:*

$$\mathbb{P}_0\left(E_x^y(n)\right) = (1 + \bar{o}(1))\frac{4}{\pi}\sin\left(\frac{\pi(x+1)}{T}\right)e^{-g(T+2)n}.$$

- *If* $\lim_{n\to\infty} \frac{n}{T^3} = \alpha \in (0, +\infty)$, *then:*

$$\mathbb{P}_0\left(E_x^y(n)\right) = (1 + \bar{o}(1))\frac{4}{\pi}(e^{\alpha\pi^2} - 1)\left[e^{\alpha\pi^2}\sin\left(\frac{\pi(x+1)}{T}\right) - \sin\left(\frac{\pi x}{T}\right)\right]e^{-g(T)n}.$$

The asymptotics may be simplified if $x = \bar{\mathcal{O}}(1)$ *(see (4.11)) or if* $x \to \infty$ *(see (4.12))*.

- *If* $\lim_{n\to\infty} \frac{n}{T^3} = 0$ *and* $n \geq \frac{1}{4}T^2\log T$, *then:*

$$\mathbb{P}_0\left(E_x^y(n)\right) = (1 + \bar{o}(1))\frac{4\pi^3 n^2}{T^6}\left[\sin\left(\frac{\pi x}{T}\right) + \frac{T^2}{\pi n}\right]e^{-g(T)n}.$$

We can summarize these results in a more compact way, if we accept to be losing some precision in the case where x is close to 0 (more precisely if $x/T \to 0$):

$$\mathbb{P}_0\left(E_x^y(n)\right) = \psi\left(\frac{n\pi^2}{T^3}\right)\left[\sin\left(\frac{\pi x}{T}\right) + \bar{o}(1)\right]e^{-g(T+2)n}, \quad \text{with } \psi(\alpha) := \frac{4}{\pi}(1 - e^{-\alpha})^2. \quad (1.4)$$

Here, the $\bar{o}(1)$ is uniform in x, y and depends only on $T = x + y$. To get (1.4), we have used in particular that $g(T+2)n - g(T)n \sim \frac{2\pi^2 n}{T^3}$ as $T \to \infty$, which converges to $2\alpha\pi^2$ if $\lim_{n\to\infty}\frac{n}{T^3} = \alpha \in [0, +\infty)$.

Comment. Whenever $x = 0$ (or $y = 0$ using symmetry) we have the same theorem applied to $x = 0$ (see Section 4.2.4), except when $\frac{n}{T^3} \to 0$ in which case

$$\mathbb{P}_0\left(E_0^T(n)\right) = (1 + \bar{o}(1))\frac{4n\pi}{(T+1)^3}\sin\left(\frac{\pi}{T+2}\right)e^{-g(T+1)n}.$$

This will not be significant starting from Section 2.2 as it only consists of two not-so-peculiar range configurations among the *many* configurations in the partition function.

Let us stress that one easily deduces from Theorem 1.4 the following statement, leading to the asymptotic independence in Theorem 1.1, as well as the convergence in distribution of $\frac{W_n}{T_n^*}$ to $\mathcal{W}$.

Proposition 1.5. *Let $(t_n)_{n\geq 1}$ be any sequence of integers such that $\lim_{n\to\infty} t_n = \infty$ and $\frac{1}{4}t_n^2 \log t_n \leq n$. Then, conditionally on $T_n = t_n$, $\frac{W_n}{t_n}$ converges in distribution to $\mathcal{W}$. More precisely, we have the following local limit convergence: uniformly for w such that $2w \in [\![-t_n, t_n]\!]$,*

$$\mathbb{P}\left(W_n = w \mid T_n = t_n\right) = \frac{\pi}{2}\left[\cos\left(\frac{w\pi}{t_n}\right) + \bar{o}(1)\right] \quad \text{as } n \to \infty.$$

Note that this proposition allows us to focus our study on T_n instead of (T_n, W_n).

Proof. For $-\frac{1}{2} \leq a \leq b \leq \frac{1}{2}$, we get thanks to (1.4) that

$$\mathbb{P}\left(a \leq \frac{W_n}{t_n} \leq b; T_n = t_n\right) = \sum_{\substack{x+y=t_n \\ 2at_n \leq y-x \leq 2bt_n}} \mathbb{P}\left(E_x^y(n)\right)$$

$$= \frac{4}{\pi}\left(1 - e^{-\frac{n\pi^2}{t_n^3}}\right)^2 e^{-g(t_n+2)n} \sum_{2at_n \leq 2w \leq 2bt_n}\left[\cos\left(\frac{w\pi}{t_n}\right) + \bar{o}(1)\right],$$

where we have set $w = w(x,y) := \frac{y-x}{2}$. Similarly,

$$\mathbb{P}\left(T_n = t_n\right) = \frac{4}{\pi}\left(1 - e^{-\frac{n\pi^2}{t_n^3}}\right)^2 e^{-g(t_n+2)n} \sum_{-t_n \leq 2w \leq t_n}\left[\cos\left(\frac{w\pi}{t_n}\right) + \bar{o}(1)\right].$$

We therefore end up with

$$\mathbb{P}\left(a \leq \frac{W_n}{t_n} \leq b \mid T_n = t_n\right) = \frac{\sum_{at_n \leq w \leq bt_n}\left[\cos\left(\frac{w\pi}{t_n}\right) + \bar{o}(1)\right]}{\sum_{-t_n \leq 2w \leq t_n}\left[\cos\left(\frac{w\pi}{t_n}\right) + \bar{o}(1)\right]} \xrightarrow[n\to\infty]{} \frac{\pi}{2}\int_a^b \cos(\pi u)\, \mathrm{d}u,$$

and taking $a = b = w/t_n$,

$$\mathbb{P}\left(W_n = w \mid T_n = t_n\right) = \frac{\cos\left(\frac{w\pi}{t_n}\right) + \bar{o}(1)}{\sum_{-t_n \leq 2w \leq t_n}\left[\cos\left(\frac{w\pi}{t_n}\right) + \bar{o}(1)\right]} = \frac{\pi}{2}\left[\cos\left(\frac{w\pi}{t_n}\right) + \bar{o}(1)\right]. \quad \square$$

1.4 Some heuristics

Let us present some heuristics for obtaining the asymptotics of the partition function, and explain how the quantities T_n^* and a_n (recall (1.2)) appear. We can decompose the partition function as

$$Z_{n,h_n} = \sum_{x,y \geq 0} e^{-h_n(T+1)} \mathbb{P}\left(E_x^y(n)\right),$$

where we have set $T = T(x,y) = x + y$. In view of Theorem 1.4, we have $\mathbb{P}\left(E_x^y(n)\right) = u_n(x,y)e^{-g(T)n}$ with $g(T) = (1 + \bar{o}(1))\frac{\pi^2}{2T^2}$. Hence, the main contribution to the sum will come from x, y with T that is close to minimizing the function

$$\varphi_n(T) := h_n T + \frac{n\pi^2}{2T^2}. \tag{1.5}$$

Then, notice that φ_n is minimal at $T = T_n^* := \left(\frac{n\pi^2}{h_n}\right)^{1/3}$, recall (1.2), and that $\varphi_n(T_n^*) = \frac{3\pi^{1/3}}{2}n^{1/3}h_n^{2/3} = \frac{3}{2}h_n T_n^*$.

Let us now factorize $e^{\varphi_n(T_n^*)}$ (and e^{h_n}) in the sum above, to get that

$$e^{\frac{3}{2}h_n T_n^*}e^{h_n} Z_{n,h_n} \approx \sum_{x,y \geq 0} u_n(x,y) \exp\left(-\left(\varphi_n(T) - \varphi_n(T_n^*)\right)\right).$$

Now, since $\varphi_n'(T_n^*) = 0$, we have $\varphi_n(T) \approx \varphi_n(T_n^*) + (T-T_n^*)^2 \varphi_n''(T_n^*)$, with $\varphi_n''(T_n^*) = \frac{3n\pi^2}{(T_n^*)^4} = \frac{1}{a_n^2}$ (recall (1.2)). In the sum above the main contribution therefore comes from values of T that are such that $\varphi_n(T) - \varphi_n(T_n^*)$ is at most of order 1, that is with $T - T_n^* = \bar{\mathcal{O}}(a_n)$.

Let us stress once more that if $\lim_{n\to\infty} n^{-1/4}h_n = 0$ then $\lim_{n\to\infty} a_n = +\infty$, whereas if $\lim_{n\to\infty} n^{-1/4}h_n = +\infty$ then $\lim_{n\to\infty} a_n = 0$. The condition $h_n \geq n^{-1/2}(\log n)^{3/2}$ ensures that $\frac{1}{4}(T_n^*)^2 \log T_n^* \leq n$ and the condition $\lim_{n\to\infty} n^{-1}h_n = 0$ ensures that $\lim_{n\to\infty} T_n^* = +\infty$.

1.5 Further comments on the results

Theorem 1.1 states that asymptotically, the polymer behaves as a random walk whose range's size T_n fluctuates around the optimal $T_n^* = \left(\frac{n\pi^2}{h_n}\right)^{1/3}$. If $h_n n^{-1/4} \to 0$ (weak penalization), then the fluctuations are Gaussian at a scale $a_n = \frac{1}{\sqrt{3}}\left(\frac{n}{h_n^4}\right)^{1/6}$. On the other hand, if $h_n n^{-1/4} \to \infty$ (strong penalization), then the fluctuations vanish and T_n is equal to either $\lfloor T_n^* \rfloor - 2$ or $\lfloor T_n^* \rfloor - 1$. In both cases, the relative position of the center of the range is asymptotically independent from its size, with distribution given by the density $\frac{\pi}{2}\cos(\pi u)\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(u)$, conjectured or discussed in previous works (see [9, Theorem 8.3] for example) but with no concrete proof (to the best of our knowledge).

1.5.1 Continuous analogue of the model

One can easily see the similarities between this polymer model and the study of the Brownian motion penalized by the amplitude of its trajectory. For a Brownian motion β, denote $|C_T| := |\{\beta_t : t \leq T\}|$ its amplitude at time T (here $|\cdot|$ is the Lebesgue measure). Donsker and Varadhan proved in [11] that

$$\lim_{T\to\infty} \frac{1}{T^{1/3}} \log \mathbb{E}\left[e^{-\nu|C_T|}\right] = -\frac{3}{2}(\nu\pi)^{2/3}.$$

Smock later expanded on this result in [16] and obtained that the associated Gibbs measures $\mathbb{P}_{T,\nu}(\mathrm{d}\omega) = e^{-\nu|C_T|}\mathbb{W}(\mathrm{d}\omega)$ (with $\mathbb{W}$ the Wiener measure) converge weakly to a measure $\mathbb{P}_{\infty,\nu}$ given by

$$\mathbb{P}_{\infty,\nu}(A) = \int_0^{c_\nu} \frac{\pi}{2c_\nu} \sin\left(\frac{\pi u}{c_\nu}\right) P_{u-c_\nu,u}(A)\,\mathrm{d}u\,,$$

with $c_\nu = (\pi^2/\nu)^{1/3}$, where $P_{u-c_\nu,u}$ denotes the path measure of a Brownian taboo process with taboo set $\{u-c_\nu, u\}$. In other words, $\mathbb{P}_{\infty,\nu}$ is a mixture of taboo processes $P_{u-c_\nu,u}$, which correspond to the actual diffusion process conditioned to stay in an interval of length c_ν and upper edge u; additionally, the mixing measure selecting the upper edge u is identical to $\mathcal{W}$ in Theorem 1.1 (if one selects the center of the range rather than the upper edge). This is therefore completely analogous to our Theorem 1.1.

However, because there is no underlying lattice, the continuous case should not display a transition for the fluctuations at $\nu = \nu_T \asymp T^{1/4}$: when $\lim_{T\to\infty} T^{-1/4}\nu_T$, fluctuations become $\bar{o}(1)$ but still remain Gaussian after a proper scaling. Let us also stress that in the continuous case, well-known results such as Lévy triple law (see [15, Theorem 6.18]) allow for relatively simple computations of the law of the endpoint β_T for a large T conditionally on the range's endpoints — which Theorem 1.4 does not provide in our setting, we only get the relative position of the starting point relative to the range, see Proposition 1.5. Obtaining a result for the starting and endpoint for our model would require the joint law of (M_n^-, M_n^+, S_n) or a study based on local times of the polymer, which are both beyond the scope of this paper.

1.5.2   Other related models

Related models for self-interacting polymers have been studied in the literature these past years. We mention here two of these models and their recent advancements.

First, one can consider a disordered version of the random walk penalized by its range, *i.e.* the case where the penalization by the range is perturbed by a random environment. Take a collection of i.i.d variables $(\omega_z)_{z\in\mathbb{Z}}$ and consider the random polymer measure

$$\mathrm{d}\mathbb{P}_{n,h}^{\omega,\beta}(S) = \frac{1}{Z_{n,h}^{\beta,\omega}} \exp\Big(\sum_{z\in\mathcal{R}_n(S)} \big(\beta\omega_z - h\big)\Big)\mathrm{d}\mathbb{P}(S),$$

in particular $\mathbb{P}_{n,h} = \mathbb{P}_{n,h}^{\omega,0}$. This quenched model was studied in [3, 4, 14], for size-dependent parameters h_n and β_n. In dimension $d = 1$, [3] finds a wide range of behaviors for the polymer depending on the sign and the growth speed of the parameters h_n, β_n. However, several questions remain open, such as determining the location and fluctuations of the range (in the spirit of Theorem 1.1) in a regime where the range size (properly rescaled) converges to a non-random quantity — we are currently investigating this question [6].

Another related model is the charged polymer, where charges are attached to the different monomers and interact with each other, see [9, Chapter 8] for an overview. Take i.i.d. random variables $(\omega_k)_{k\in\mathbb{N}}$, and consider the following quenched Gibbs measure on random walk trajectories

$$\mathrm{d}\mathbb{P}_{n,\beta}^{\omega}(S) = \frac{1}{Z_{n,\beta}^{\omega}} \exp\Big(-\beta \sum_{1\le i<j\le n} \omega_i\omega_j \mathbb{1}_{\{S_i=S_j\}}\Big)\mathrm{d}\mathbb{P}(S)\,.$$

Some recent papers [2, 7] are dealing with the annealed version of the model, that can be written in the following form

$$d\mathbb{P}_{n,\beta}^{\mathrm{ann}}(S) = \frac{1}{Z_{n,\beta}^{\mathrm{ann}}} \exp\Big(-\sum_{x\in\mathbb{Z}^d} g_\beta(\ell_n(x))\Big) d\mathbb{P}(S),$$

where $\ell_n(x) = \sum_{i=1}^n \mathbb{1}_{\{S_i=x\}}$ is the local time at site x and where g_β is a function that depends on β and on the distribution of ω. This model has been shown to undergo a folding/unfolding phase transition, and the case of dimension $d=1$ has been investigated in remarkable detail in [7]. Our model falls in the same class of models, it corresponds to using the function $h\mathbb{1}_{\{\ell_n(x)>0\}}$ instead of the function $g_\beta(\ell_n(x))$; note that our model also displays a folding/unfolding transition when h goes from positive to negative values.

Organization of the rest of the paper

The rest of the paper is organized as follows:

- In Section 2 we focus on the case of a "weak" penalization, that is $\lim_{n\to\infty} n^{-1/4}h_n = 0$: we give local asymptotic estimates for the partition function (Lemma 2.1), from which we deduce the first point of both Theorem 1.2 and Theorem 1.1 (in that order).

- In Section 3 we treat the case of a "strong" penalization, that is $\liminf_{n\to\infty} n^{-1/4}h_n > 0$: we modify the arguments of Section 2 to provide local asymptotic estimates for the partition function (Lemma 2.1). From this, we deduce first the second point of Theorems 1.1-1.2, *i.e.* in the case $\lim_{n\to\infty} n^{-1/4}h_n = +\infty$, before we turn to the border case of Proposition 1.3, *i.e.* $\lim n^{-1/4}h_n = \hat{h} \in (0,+\infty)$.

- Finally, in Section 4 we derive sharp gambler's ruin estimates (see Lemmas 4.2-4.3) and their consequences for the range of a random walk, that is we prove Theorem 1.4.

2 Weak penalization: the case $\lim_{n\to\infty} n^{-1/4}h_n = 0$

2.1 Local asymptotics for the partition function

Our first preliminary result computes the contribution of the partition function from trajectories with a fixed size of the range T_n, with $T_n = T_n^* + \bar{o}(T_n^*)$. Recall that $T_n^* := \left(\frac{n\pi^2}{h_n}\right)^{1/3}$ and $a_n := \frac{1}{\sqrt{3n\pi^2}}(T_n^*)^2$.

Lemma 2.1. *Assume that $h_n \geq n^{-1/2}(\log n)^{3/2}$ and that $\lim_{n\to\infty} n^{1/4}h_n = 0$. Let $(\varepsilon_n)_{n\geq 1}$ be any vanishing sequence. Then, for any $t \in \mathbb{Z}$ and $w \in \frac{1}{2}\mathbb{Z}$ such that $|t| \leq \varepsilon_n T_n^*$ and $w - \frac{1}{2}(\lfloor T_n^* \rfloor + t) \in 2\mathbb{Z}$, we have*

$$Z_{n,h_n}\Big(T_n = \lfloor T_n^* \rfloor + t, W_n = w\Big) = \psi_n \times \Big(\cos\Big(\frac{\pi w}{T_n^*}\Big) + \bar{o}(1)\Big) \times e^{-(1+\bar{o}(1))\frac{t^2}{2a_n^2}}, \qquad (2.1)$$

where the $\bar{o}(1)$ only depends on ε_n, and where we have set:

$$\psi_n := \psi(h_n)\exp\Big(-h_n(T_n^*+1) - g(T_n^*+2)n\Big), \qquad \text{with } \psi(\alpha) = \frac{4}{\pi}(1 - e^{-\alpha})^2. \qquad (2.2)$$

8

Proof. We have

$$Z_{n,h_n}\left(T_n = \lfloor T_n^* \rfloor + t, W_n = w\right) = e^{-h_n(x+y+1)}\mathbb{P}\left(E_x^y(n)\right),$$

with $x + y = \lfloor T_n^* \rfloor + t$ and $\frac{1}{2}(y - x) = w$. Thanks to Theorem 1.4, we can estimate this term. Indeed, for every x, y such that $\lim_{n\to\infty} \frac{x+y}{T_n^*} = 1$, using (1.4) we have

$$\mathbb{P}\left(E_x^y(n)\right) = \psi\left(\frac{n\pi^2}{(T_n^*)^3}\right)\left(\sin\left(\frac{\pi x}{T_n^*}\right) + \bar{o}(1)\right)e^{-g(x+y+2)n}, \tag{2.3}$$

where the $\bar{o}(1)$ is uniform in x, y and $\psi(\alpha) = \frac{4}{\pi}(1 - e^{-\alpha\pi^2})^2$. Note that by definition of T_n^* we have $\frac{n\pi^2}{(T_n^*)^3} = h_n$.

Now, here we have that $x = \frac{1}{2}(\lfloor T_n^* \rfloor + t) - w$, with $\frac{1}{2T_n^*}(\lfloor T_n^* \rfloor + t) \to \frac{1}{2}$. Hence, we can write $\sin(\frac{\pi x}{T_n^*}) = \cos(\frac{\pi w}{T_n^*}) + \bar{o}(1)$ in (2.3). We also write $h_n(T + 1) + g(T + 2)n = \varphi_n(T) + h_n + \tilde{g}(T)n$, with $\tilde{g}(T) = g(T + 2) - \frac{\pi^2}{2T}$, to get that

$$
\begin{aligned}
&Z_{n,h_n}\left(T_n = \lfloor T_n^* \rfloor + t, W_n = w\right) \\
&= \psi(h_n)\left(\cos\left(\frac{\pi w}{T_n^*}\right) + \bar{o}(1)\right)\exp\left(-\varphi_n(\lfloor T_n^* \rfloor + t) - h_n - \tilde{g}(\lfloor T_n^* \rfloor + t)n\right).
\end{aligned}
\tag{2.4}
$$

We can use that $\varphi_n'(T_n^*) = 0$, $\varphi_n''(T_n^*) = \frac{3n\pi^2}{(T_n^*)^4} = 1/a_n^2$ and $\varphi_n'''(T) = -\frac{12n\pi^2}{T^5}$ to get that for any $\frac{1}{2}T_n^* \leq T \leq 2T_n^*$

$$\left|\varphi_n(T) - \varphi_n(T_n^*) - \frac{(T - T_n^*)^2}{2a_n^2}\right| \leq C\frac{|T - T_n^*|^3 n}{(T_n^*)^5} = C'\frac{|T - T_n^*|^3}{T_n^* a_n^2}.$$

Hence, using that $a_n \to \infty$, we get that

$$\varphi_n(\lfloor T_n^* \rfloor + t) = \varphi_n(T_n^*) - \frac{t^2}{2a_n^2} + \bar{o}\left(\frac{1}{a_n^2}\right). \tag{2.5}$$

We can also perform the same expansion for $\tilde{g}(T) = g(T + 2) - \frac{\pi^2}{2T}$, for which $\tilde{g}'(T) = \frac{\pi^2}{T^3} - \frac{\pi}{(T+2)^2}\tan(\frac{\pi}{T+2})$:

$$|\tilde{g}(T + 2) - \tilde{g}(T_n^*)| \leq C\frac{|T + 2 - T_n^*|}{(T_n^*)^4},$$

so that, inserting $T = \lfloor T_n^* \rfloor + t$

$$|\tilde{g}(\lfloor T_n^* \rfloor + t + 2)n - \tilde{g}(T_n^*)n| \leq C\frac{|t + \delta_n + 2|}{a_n^2} = \bar{o}(1)\frac{t^2}{a_n^2} + \bar{o}(1) \tag{2.6}$$

as $n \to \infty$, uniformly in t (consider separately the case $|t| \leq a_n$ and $|t| \geq a_n$).

All together, plugging (2.5)-(2.6) into (2.4), we end up with the desired result, with $\psi_n := \psi(h_n)\exp(-h_n - \varphi_n(T_n^*) - \tilde{g}(T_n^*)n)$ which coincides with the definition (2.2) above. $\qquad\square$

2.2 Asymptotics of the partition function

Lemma 2.1 allows us to obtain the correct behavior for the partition function.

Proof of Theorem 1.2. Assume that $h_n \geq n^{-1/2}(\log n)^{3/2}$ and that $\lim_{n\to\infty} n^{-1/4}h_n = 0$, so in particular $a_n \to +\infty$.

Note that by [3] (or simply using large deviation principles), we have for any $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}_{n,h_n}\left(|T_n - T_n^*| > \varepsilon T_n^*\right) = 0\,.$$

Therefore, one can find some vanishing sequence $(\varepsilon_n)_{n\geq 0}$ such that we have the asymptotic equivalence $Z_{n,h_n} = (1 + \bar{o}(1))Z_{n,h_n}(|T_n - T_n^*| \leq \varepsilon_n T_n^*)$. We therefore only have to estimate that last partition function. We may decompose it as

$$Z_{n,h_n}\left(|T_n - T_n^*| \leq \varepsilon_n T_n^*\right) = \sum_{t=\lfloor T_n^*\rfloor - \lfloor\varepsilon_n T_n^*\rfloor}^{\lfloor T_n^*\rfloor + \lfloor\varepsilon_n T_n^*\rfloor} \sum_{-(T_n^*+t)\leq 2w \leq T_n^*+t} Z_{n,h_n}\left(T_n = \lfloor T_n^*\rfloor + t\,, W_n = w\right)\,.$$

Therefore, thanks to Lemma 2.1, using that $\frac{|t|}{a_n^2} = \bar{o}(1)\frac{t^2}{a_n^2} + \bar{o}(1)$ as $n \to \infty$ with a $\bar{o}(1)$ uniform in t (since $a_n \to \infty$), we get that

$$Z_{n,h_n}\left(|T_n - T_n^*| \leq \varepsilon_n T_n^*\right) = (1 + \bar{o}(1))\psi_n \sum_{t=-\lfloor\varepsilon_n T_n^*\rfloor}^{\lfloor\varepsilon_n T_n^*\rfloor} e^{-(1+\bar{o}(1))\frac{t^2}{2a_n^2}} \sum_{\substack{-(T_n^*+t)\leq 2w \leq T_n^*+t \\ w \in \frac{1}{2}(\lfloor T_n^*\rfloor+t)+\mathbb{Z}}} \cos\left(\frac{\pi w}{T_n^*}\right)\,.$$

Now, as T_n^* goes to $+\infty$, the internal sum is a Riemann sum: we have, uniformly for $(1 - \varepsilon_n)T_n^* \leq t \leq (1 + \varepsilon_n)T_n^*$,

$$\sum_{\substack{-(T_n^*+t)\leq 2w \leq T_n^*+t \\ w \in \frac{1}{2}(\lfloor T_n^*\rfloor+t)+\mathbb{Z}}} \cos\left(\frac{\pi w}{T_n^*}\right) \sim T_n^* \int_{-\frac{1}{2}}^{\frac{1}{2}} \cos(\pi v)\,\mathrm{d}v = \frac{2}{\pi}T_n^*\,.$$

Then, a sum over t remains, which is also a Riemann sum: as $a_n \to +\infty$ and $\varepsilon_n T_n^*/a_n \to +\infty$, we have

$$\sum_{t=-\lfloor\varepsilon_n T_n^*\rfloor}^{\lfloor\varepsilon_n T_n^*\rfloor} e^{-(1+\bar{o}(1))\frac{t^2}{2a_n^2}} \sim a_n \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}}\,\mathrm{d}u = \sqrt{2\pi}a_n\,.$$

All together, we have proved that, as $n \to \infty$

$$Z_{n,h_n} \sim \frac{2\sqrt{2}}{\sqrt{\pi}}\psi_n a_n T_n^*\,. \tag{2.7}$$

Recalling the definition of T_n^* and a_n, we have $a_n T_n^* = \frac{\pi}{\sqrt{3}}\frac{\sqrt{n}}{h_n}$. Additionally, recalling the definition (2.2) of ψ_n, and using that $g(T + 2) = \frac{\pi^2}{2T^2} - \frac{2\pi^2}{T^3} + \bar{\mathcal{O}}(\frac{1}{T^4})$ as $T \to \infty$, we get that

$$\psi_n = \psi(h_n)\exp\left(-h_n T_n^* - h_n - \frac{n\pi^2}{2(T_n^*)^2} + \frac{2\pi^2 n}{(T_n^*)^3} + \bar{o}(1)\right)\,,$$

since $\lim_{n\to\infty}\frac{n}{(T_n^*)^4} = 0$ because $\lim_{n\to\infty} n^{-1/4}h_n = 0$. Using that by definition of T_n^* we have $\frac{n\pi^2}{(T_n^*)^3} = h_n$, we get that $\psi_n \sim \psi(h_n)e^{h_n}e^{-\frac{3}{2}h_n T_n^*}$. Putting all together and noting that $e^\alpha(1 - e^{-\alpha})^2 = 2(\cosh(\alpha) - 1)$, this concludes the proof. $\square$

Proof of Theorem 1.1. The proof reduces to showing the following Lemma.

Lemma 2.2. *Let $h_n \geq n^{-1/2}(\log n)^{3/2}$ be such that $\lim\limits_{n \to \infty} n^{-1/4}h_n = 0$. Then, for any $r < s$, we have*

$$\lim_{n \to \infty} \frac{1}{\psi_n a_n T_n^*} Z_{n,h_n}\left(r \leq \frac{|T_n - T_n^*|}{a_n} \leq s\right) = \frac{2}{\pi} \int_r^s e^{-\frac{u^2}{2}}\,\mathrm{d}u\,,$$

where ψ_n is the sequence that appears in Lemma 2.1.

Indeed, once we have this lemma, in view of the asymptotics (2.7) and Proposition 1.5, we get that for any $r < s$ and any $a < b$,

$$\mathbb{P}_{n,h_n}\left(r \leq \frac{|T_n - T_n^*|}{a_n} \leq s, a \leq \frac{W_n}{T_n^*} \leq b\right) = \frac{1}{Z_{n,h_n}} Z_{n,h_n}\left(r \leq \frac{|T_n - T_n^*|}{a_n} \leq s, a \leq \frac{W_n}{T_n^*} \leq b\right)$$

$$\xrightarrow{n \to \infty} \int_r^s \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\,\mathrm{d}u \int_a^b \frac{\pi}{2}\cos(\pi v)\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}\mathrm{d}v\,,$$

which concludes the proof. $\qquad\qquad\square$

Proof of Lemma 2.2. The proof proceeds as for the proof of Theorem 1.2. We can decompose the partition function as

$$\sum_{t=\lfloor ra_n\rfloor}^{\lfloor sa_n\rfloor} \sum_{\substack{\lfloor 2aT_n^*\rfloor \leq 2w \leq \lfloor 2bT_n^*\rfloor \\ |2w|\leq \lfloor T_n^*\rfloor+t}} Z_{n,h_n}\left(T_n = \lfloor T_n^*\rfloor + t\,, W_n = w\right)$$

$$= (1 + \bar{o}(1))\psi_n \sum_{t=\lfloor ra_n\rfloor}^{\lfloor sa_n\rfloor} e^{-(1+\bar{o}(1))\frac{t^2}{2a_n^2}} \sum_{\substack{-\lfloor T_n^*\rfloor-t\leq 2w\leq \lfloor T_n^*\rfloor+t \\ w\in \frac{1}{2}(\lfloor T_n^*\rfloor+t)+\mathbb{Z}}} \cos\left(\frac{\pi w}{T_n^*}\right)\,,$$

where we have used Lemma 2.1 as above (using that $a_n \to \infty$).

Again, as T_n^* goes to $+\infty$, the internal sum is a Riemann sum: we have, uniformly for $\lfloor T_n^*\rfloor + \lfloor sa_n\rfloor \leq t \leq \lfloor T_n^*\rfloor + \lfloor sa_n\rfloor$,

$$\sum_{\substack{-\lfloor T_n^*\rfloor-t\leq 2w\leq \lfloor T_n^*\rfloor+t \\ w\in \frac{1}{2}(\lfloor T_n^*\rfloor+t)+\mathbb{Z}}} \cos\left(\frac{\pi w}{T_n^*}\right) \sim T_n^* \int_{-\frac{1}{2}}^{\frac{1}{2}} \cos(\pi v)\,\mathrm{d}v \sim \frac{2}{\pi} T_n^*\,.$$

Then, the sum over t that remains is also a Riemann sum: as $a_n \to +\infty$, we have

$$\sum_{t=\lfloor sa_n\rfloor}^{\lfloor ta_n\rfloor} e^{-(1+\bar{o}(1))\frac{t^2}{2a_n^2}} \sim a_n \int_s^t e^{-\frac{u^2}{2}}\,\mathrm{d}u\,,$$

which concludes the proof. $\qquad\qquad\square$

3 Strong penalization and vanishing fluctuations

3.1 The case $\liminf\limits_{n\to\infty} n^{-1/4} h_n = +\infty$

See that the case where $a_n \to 0$ is much more restrictive to establish an analog of Lemma 2.1, as $\bar{\mathcal{O}}(a_n^{-2})$ quantities now bring extremely large contributions to the exponential part of Z_{n,h_n} and slight deviations from the optimal size T_n^* will be penalized by a large factor. Indeed, if we are to get a *continuity* from Theorem 1.1 when $\limsup_{n\to\infty} a_n < \infty$, we want to know the exact asymptotic law of fluctuations without renormalization. We denote

$$\bar{\varphi}_n(T) = h_n(T+1) + \frac{n\pi^2}{2(T+2)^2} \tag{3.1}$$

and $T_n^o := \arg\min \bar{\varphi}_n(T)$ the (more exact) optimal amplitude of the range.

Lemma 3.1. *Assume that $\lim_{n\to\infty} n^{-1/4} h_n = +\infty$ and $\lim_{n\to\infty} n^{-1} h_n = 0$ and let $(\varepsilon_n)_{n\geq 1}$ be any vanishing sequence. Then, for any $t \in \mathbb{Z}$ and $w \in \frac{1}{2}\mathbb{Z}$ such that $|t| \leq \varepsilon_n T_n^o$ and $w - \frac{1}{2}(\lfloor T_n^o \rfloor + t) \in 2\mathbb{Z}$, we have*

$$Z_{n,h_n}\left(T_n = \lfloor T_n^* - 2 \rfloor + t, W_n = w\right) = \bar{\psi}_n \times \left(\cos\left(\frac{\pi w}{T_n^*}\right) + \bar{o}(1) \right) \times e^{-(1+\bar{o}(1))\frac{(t-t_n^o)^2}{2a_n^2}}, \tag{3.2}$$

where $t_n^o := T_n^o - \lfloor T_n^o \rfloor$, $\bar{o}(1)$ is a vanishing quantity that depends only on ε_n and

$$\bar{\psi}_n := \psi(h_n) \exp\left(-h_n(T_n^* - 1) - g(T_n^*)n \right), \quad \text{with} \quad \psi(\alpha) = \frac{4}{\pi}(1 - e^{-\alpha})^2. \tag{3.3}$$

Proof. We can perform the same decomposition as in Lemma 2.1 and setting $T = \lfloor T_n^* - 2 \rfloor + t$, we arrive at (analogously to (2.4))

$$\begin{aligned}
Z_{n,h_n}\left(T_n = T, W_n = w\right) &= \psi(h_n)\left(\cos\left(\frac{\pi w}{T}\right) + \bar{o}(1) \right) \times e^{-h_n(T+1) - g(T+2)n} \\
&= \psi(h_n)\left(\cos\left(\frac{\pi w}{T}\right) + \bar{o}(1) \right) \times e^{-\bar{\varphi}_n(T) - \bar{g}(T)n}
\end{aligned} \tag{3.4}$$

with $\bar{\varphi}_n$ defined above in (3.1) and where $\bar{g}(T) := \frac{\pi^2}{T+2} - g(T+2)$. One can then easily check that $T_n^o = T_n^* - 2$ and that $\bar{\varphi}_n''(T_n^o) = -\frac{3n\pi^2}{(T_n^o+2)^4}$ and $\bar{\varphi}_n^{(3)}(T_n^o) = \frac{12n\pi^2}{(T_n^o+2)^5}$, thus

$$\left| \bar{\varphi}_n(T) - \bar{\varphi}_n(T_n^o) - \frac{3n\pi^2}{2(T_n^o+2)^4}(T - T_n^o)^2 \right| \leq \frac{12n\pi^2}{(T_n^o+2)^5}(T - T_n^o)^3 = \bar{o}(1)\frac{(T - T_n^o)^3}{a_n^2}.$$

Furthermore, the first two orders of the Taylor expansion of $\bar{g}(T)$ around T_n^o are given by

$$\begin{aligned}
\bar{g}(T_n^o) + (T - T_n^o) &\left(\frac{\pi^2}{(T_n^o+2)^3} - \frac{\pi}{(T_n^o+2)^2}\tan\frac{\pi}{T_n^o+2} \right) \\
+ \frac{(T - T_n^o)^2}{2} &\left(-\frac{3\pi^2}{(T_n^o+2)^4} + \frac{2\pi}{(T_n^o+2)^3}\tan\frac{\pi}{T_n^o+2} + \frac{\pi^2}{(T_n^o+2)^4}\left(1 + \tan^2\frac{\pi}{T_n^o+2}\right) \right).
\end{aligned}$$

A Taylor expansion of the tangent leads to the following bound

$$\begin{aligned}
n|\bar{g}(T) - \bar{g}(T_n^o)| \\
\leq |T - T_n^o|\frac{n\pi^3}{3(T_n^o+2)^5} + (T - T_n^o)^2\frac{n\pi^3}{3(T_n^o+2)^6} = \frac{1}{a_n^2}\bar{o}(1)\left[|T - T_n^o| + (T - T_n^o)^2 \right].
\end{aligned}$$

In particular, we get the lemma injecting $T = \lfloor T_n^* - 2 \rfloor + t = T_n^o - t_n^o + t$ in (3.4), using also that $\bar{\varphi}_n(T_n^o) + n\bar{g}(T_n^o) = h_n(T_n^* - 1) + ng(T_n^*)$. $\qquad\square$

Proof of Theorem 1.2. Suppose $\liminf_{n\to\infty} n^{-1/4}h_n = +\infty$ then $a_n \to 0$ and we can't apply Riemann summations as in the proof of Theorem 1.2 or Lemma 2.2. However Lemma 3.1 allows to exclude the slightest deviations using $a_n^{-2} \to \infty$ that we use to estimate

$$Z_{n,h_n}\left(|T_n - T_n^*| \leq \varepsilon_n T_n^*\right) = (1 + \bar{o}(1))\psi_n \sum_{t=-\lfloor \varepsilon_n T_n^* \rfloor}^{\lfloor \varepsilon_n T_n^* \rfloor} e^{-(1+\bar{o}(1))\frac{(t-t_n^o)^2}{2a_n^2}} \sum_{\substack{-(T_n^*+t)\leq 2w \leq T_n^*+t \\ w \in \frac{1}{2}(\lfloor T_n^* \rfloor + t) + \mathbb{Z}}} \cos\left(\frac{\pi w}{T_n^*}\right).$$

As T_n^* goes to $+\infty$, the internal sum still is a Riemann sum and thus, uniformly for $(1 - \varepsilon_n)T_n^o \leq t \leq (1 + \varepsilon_n)T_n^o$,

$$\sum_{\substack{-(T_n^*+t)\leq 2w \leq T_n^*+t \\ w \in \frac{1}{2}(\lfloor T_n^* \rfloor + t) + \mathbb{Z}}} \cos\left(\frac{\pi w}{T_n^*}\right) \sim T_n^* \int_{-\frac{1}{2}}^{\frac{1}{2}} \cos(\pi v)\,\mathrm{d}v = \frac{2}{\pi} T_n^*.$$

Now, for the sum that remains, we factorize by the largest term, attained at $t = 0$ or $t = 1$ depending on t_n^o. Thus denote $\delta_n = t - t_n^o$ with $t = 0$ or 1 such that $\delta_n^2 = (-t_n^o)^2 \wedge (1 - t_n^o)^2$ meaning $\delta_n = \mathbb{1}_{\{t_n^o \geq \frac{1}{2}\}} - t_n^o$. We have

$$\sum_{t=-\lfloor \varepsilon_n T_n^* \rfloor}^{\lfloor \varepsilon_n T_n^* \rfloor} e^{-\frac{(t-t_n^o)^2}{2a_n^2} + \bar{o}(a_n^{-2})} = e^{-(1+\bar{o}(1))\frac{\delta_n^2}{2a_n^2}}\left(1 + e^{-\frac{(1-\delta_n-2t_n^o)^2-\delta_n^2}{2a_n^2}} + \sum_{\substack{t=-\lfloor \varepsilon_n T_n^* \rfloor \\ t \neq 0,1}}^{\lfloor \varepsilon_n T_n^* \rfloor} e^{\frac{\delta_n^2-(t-t_n^o)^2}{2a_n^2}}\right). \quad (3.5)$$

First use dominated convergence to get

$$0 \leq \sum_{\substack{t=-\lfloor \varepsilon_n T_n^* \rfloor \\ t \neq 0,1}}^{\lfloor \varepsilon_n T_n^* \rfloor} e^{\frac{\delta_n^2-(t-t_n^o)^2}{2a_n^2}} \leq \sum_{\substack{t \in \mathbb{Z} \\ t \neq 0,1}} e^{-\frac{1}{2a_n^2}(t-1)(t+1-t_n^o)} \xrightarrow[n\to\infty]{} 0.$$

Then, note that when $t_n^o \neq \frac{1}{2}$, the second term of (3.5) goes to 0, whereas when $t_n^o = \frac{1}{2}$ it is equal to 1. Thus we have

$$Z_{n,h_n} = \frac{2}{\pi}\left(1 + \mathbb{1}_{\{t_n^o = \frac{1}{2}\}}\right)\psi(h_n)e^{h_n - \frac{3}{2}h_n T_n^* - \frac{\delta_n^2}{2a_n^2} + \bar{o}(a_n^{-2})} \exp\left(g(T_n^* + 2)n - \frac{n\pi^2}{2T_n^*} - \frac{2}{(T_n^*)^3}\right),$$

where we recall the definition $\psi(a) = \frac{4}{\pi}(1 - e^{-a})^2$. Note that $g(T_n^* + 2)n - \frac{n\pi^2}{2T_n^*} - \frac{2}{(T_n^*)^3} = \bar{\mathcal{O}}\left(\frac{n}{(T_n^*)^4}\right) = \bar{\mathcal{O}}(a_n^{-2})$ now goes to infinity. We could push the asymptotic expansion to any arbitrary order, but since we already have the term $\bar{o}(a_n^{-2}) = \bar{o}\left(\frac{n}{(T_n^*)^4}\right)$, we simply use the following expansion up to order 4:

$$g(T_n^* + 2)n - \frac{n\pi^2}{2T_n^*} - \frac{2}{(T_n^*)^3} = (72 + \pi^2)\frac{n\pi^2}{12(T_n^*)^4}(1 + \bar{o}(1)).$$

Using $e^\alpha(1 - e^{-\alpha})^2 = 2(\cosh(\alpha) - 1)$ and writing a_n^{-2} explicitly concludes the proof. $\quad\square$

Proof of Theorem 1.1. The above proof of Theorem 1.2 already shows that

$$\frac{1}{Z_{n,h_n}}Z_{n,h_n}\left(|T_n - \lfloor T_n^* - 2 \rfloor| \in \{0,1\}\right) \xrightarrow[n\to\infty]{} 0.$$

According to (3.5), the proof only consists of finding which values of t contribute to $e^{-\frac{\delta_n^2}{a_n^2}}$. By the definition of δ_n we easily see that if $t_n^o < \frac{1}{2}$ it is $t = 0$, if $t_n^o > \frac{1}{2}$ it is $t = 1$, and when $t_n^o = \frac{1}{2}$ both have the same contribution. This observation leads to defining $\mathcal{A}_n$ as announced in Theorem 1.1, and Proposition 1.5 completes the proof. $\quad\square$

3.2 Case $\lim_{n\to\infty} n^{-1/4}h_n = \hat{h}$: order one fluctuations

Proof of Proposition 1.3. Going back to using Lemma 3.1, we get for any vanishing sequence $(\varepsilon_n)_{n\geq 1}$

$$Z_{n,h_n}\left(|T_n-T_n^*| \leq \varepsilon_n T_n^*\right) = (1+\bar{o}(1))\bar{\psi}_n \sum_{t=-\lfloor \varepsilon_n T_n^*\rfloor+2}^{\lfloor \varepsilon_n T_n^*\rfloor+2} e^{-(1+\bar{o}(1))\frac{(t-t_n^o)^2}{2a_n^2}} \sum_{\substack{-(T_n^*+t)\leq 2w \leq T_n^*+t \\ w\in\frac{1}{2}(\lfloor T_n^*\rfloor+t)+\mathbb{Z}}} \cos\left(\frac{\pi w}{T_n^*}\right).$$

The internal Riemann sum is dealt with the same method as before, while we can take the limit for a_n in the external sum. Thus we have, as $n \to \infty$

$$Z_{n,h_n}\left(|T_n-T_n^*| \leq \varepsilon_n T_n^*\right) \sim \frac{2}{\pi}T_n^*\bar{\psi}_n \sum_{t=-\lfloor \varepsilon_n T_n^*\rfloor}^{\lfloor \varepsilon_n T_n^*\rfloor} e^{-\frac{(t-t_n^o)^2}{2a^2}},$$

which clearly gives that $Z_{n,h_n} \sim \frac{2}{\pi}\bar{\psi}_n T_n^*\theta_n(a)$, because we already know that taking ε_n going to zero sufficienly slowly we have $\lim_{n\to\infty}\mathbb{P}_{n,h_n}(|T_n - T_n^*| > \varepsilon_n T_n) = 0$.

Moreover, applying again Lemma 3.1, we also get that for any fixed integer $t \in \mathbb{Z}$,

$$Z_{n,h_n}\left(T_n = \lfloor T_n^* - 2\rfloor + t\right) \sim \frac{2}{\pi}T_n^*\bar{\psi}_n e^{-\frac{(t-t_n^o)^2}{2a^2}},$$

using the same calculation as above. This concludes the proof of Proposition 1.3. $\qquad\square$

4 Range endpoints and gambler's ruin estimates

4.1 Gambler's ruin estimates

We consider a band $[0,T]$ with T some positive integer, and choose a starting point $0 \leq z \leq T$. We denote $\tau_0 := \min\{n \geq 0, S_n = 0\}$, resp. $\tau_T := \min\{n \geq 0, S_n = T\}$, the hitting time of the edge at 0, resp. at T. We also denote $\tau := \tau_0 \wedge \tau_T$. We recall the formulae of [13, §14.5] for the ruin problem, in the case of a symmetric walk. We use the notation $n \leftrightarrow z$ if $n - z$ is even.

Proposition 4.1. *For any $z \in [\![0, T]\!]$ and $n > 1$,*

$$\mathbb{P}_z\left(\tau = \tau_0 = n\right) = \frac{2}{T}\sum_{1\leq k<T/2}\cos^{n-1}\left(\frac{\pi k}{T}\right)\sin\left(\frac{\pi k z}{T}\right)\sin\left(\frac{\pi k}{T}\right)\mathbb{1}_{\{n\leftrightarrow z\}}. \qquad (4.1)$$

By symmetry, we also have $\mathbb{P}_z\left(\tau = \tau_T = n\right) = \mathbb{P}_{T-z}\left(\tau = \tau_0 = n\right)$:

$$\mathbb{P}_z\left(\tau = \tau_T = n\right) = \frac{2}{T}\sum_{1\leq k<T/2}(-1)^{k+1}\cos^{n-1}\left(\frac{\pi k}{T}\right)\sin\left(\frac{\pi k z}{T}\right)\sin\left(\frac{\pi k}{T}\right)\mathbb{1}_{\{n\leftrightarrow T-z\}}.$$

$$(4.2)$$

Note that if $T - z \leftrightarrow n \leftrightarrow z$, we have

$$\mathbb{P}_z\left(\tau = n\right) = \frac{4}{T}\sum_{1\leq k<T/4}\cos^{n-1}\left(\frac{(2k-1)\pi}{T}\right)\sin\left(\frac{(2k-1)\pi z}{T}\right)\sin\left(\frac{(2k-1)\pi}{T}\right).$$

Let us now give the *sharp* asymptotic behavior of the probabilities (4.1)-(4.2) above. Recall the definition (1.3): $g(T) = -\log\cos(\frac{\pi}{T})$. By symmetry, we only deal with the case $z \in [\![0, \frac{T}{2}]\!]$.

Lemma 4.2. *Suppose that $T \to \infty$ and that $\lim_{n\to\infty} \frac{n}{T^2} = +\infty$. Then, we have the following asymptotics: for all $z \in [\![0, \frac{T}{2}]\!]$,*

$$\mathbb{P}_z \left(\tau = \tau_0 = n \right) = \left(1 + \bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}}) \right) \frac{2}{T} \sin\left(\frac{z\pi}{T} \right) \tan\left(\frac{\pi}{T} \right) e^{-g(T)n} \mathbb{1}_{\{n \leftrightarrow z\}}, \qquad (4.3)$$

$$\mathbb{P}_z \left(\tau = \tau_T = n \right) = \left(1 + \bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}}) \right) \frac{2}{T} \sin\left(\frac{z\pi}{T} \right) \tan\left(\frac{\pi}{T} \right) e^{-g(T)n} \mathbb{1}_{\{n \leftrightarrow T-z\}}. \qquad (4.4)$$

Here, $\bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}})$ is uniform in z.

Proof. The proof is inspired by [8, Appendix B], but we need here a slightly sharper version. In (4.1) and (4.2) we denote $V_0 = V_0(n, T)$ the first term:

$$V_0 = \frac{2}{T} \cos^{n-1}\left(\frac{\pi}{T} \right) \sin\left(\frac{\pi z}{T} \right) \sin\left(\frac{\pi}{T} \right) = \frac{2}{T} \sin\left(\frac{z\pi}{T} \right) \tan\left(\frac{\pi}{T} \right) e^{-g(T)n}.$$

It remains to control the remaining terms. We let

$$V_1 := \frac{2}{T} \sum_{2 \leq k < T/2} \cos^{n-1}\left(\frac{\pi k}{T} \right) \sin\left(\frac{\pi k z}{T} \right) \sin\left(\frac{\pi k}{T} \right),$$

and we only need to bound V_1/V_0. Using the bounds $\frac{2}{\pi} x \leq \sin(x) \leq x$ for $x \in [0, \frac{\pi}{2}]$, we get that

$$\frac{V_1}{V_0} \leq \frac{\pi^2}{4} \sum_{2 \leq k < T/2} k^2 \left(\frac{\cos\left(\frac{\pi k}{T} \right)}{\cos\left(\frac{\pi}{T} \right)} \right)^{n-1}.$$

Now, as $\frac{k}{T} \to 0$, we have

$$\frac{\cos\left(\frac{\pi k}{T} \right)}{\cos\left(\frac{\pi}{T} \right)} = 1 - \frac{\pi^2(k^2 - 1)}{2T^2} \left(1 + \bar{\mathcal{O}}\left(\frac{k^2}{T^2} \right) \right).$$

Hence, the l.h.s. is bounded by $\exp(-\frac{2\pi^2(k^2-1)}{5T^2})$ provided that $\frac{k}{T} \leq \varepsilon$, for some given $\varepsilon \in (0, \frac{1}{2})$. If $\frac{k}{T} \geq \varepsilon$, we can simply bound $\cos\left(\frac{\pi k}{T} \right) \leq \cos(\pi\varepsilon) \leq e^{-\frac{1}{2}\pi^2\varepsilon^2}$. We therefore get that V_1/V_0 is bounded by a constant times

$$\sum_{2 \leq k < \varepsilon T} k^2 e^{-\frac{2n\pi^2(k^2-1)}{5T^2}} + \sum_{\varepsilon T \leq k < T/2} k^2 e^{-\frac{1}{2}n\pi^2\varepsilon^2}.$$

For the first sum, we write that it is

$$e^{\frac{2\pi^2 n}{5T^2}} T^3 \times \frac{1}{T} \sum_{2 \leq k < \varepsilon T} \frac{k^2}{T^2} e^{-\frac{2\pi^2 n k^2}{5T^2}} \leq e^{\frac{2\pi^2 n}{5T^2}} T^3 \int_{2/T}^{\infty} x^2 e^{-\frac{2\pi^2}{5} n x^2} dx$$

$$= e^{\frac{2\pi^2 n}{5T^2}} \frac{T^3}{n^{3/2}} \int_{2\sqrt{n}/T}^{\infty} u^2 e^{-\frac{2\pi^2}{5} u^2} dx \leq C \frac{T^2}{n} \exp\left(-\frac{7\pi^2 n}{5T^2} \right),$$

using that $\int_v^{\infty} u^2 e^{-\frac{2\pi^2}{5} u^2} du \sim (cst.) v e^{-\frac{2\pi^2}{5} v^2}$ as $v \to \infty$. This term is therefore bounded by a constant times $\exp(-\frac{\pi^2 n}{T^2})$, as n/T^2 goes to infinity.

For the other sum, we bound it by a constant times

$$T^3 \exp\left(-\frac{1}{2} n\pi^2\varepsilon^2 \right) \leq n^{3/2} \exp\left(-\frac{1}{2} n\pi^2\varepsilon^2 \right) = \bar{o}\left(\exp\left(-\frac{\pi^2 n}{T^2} \right) \right),$$

as $\frac{n}{T^2} \to +\infty$ and $T \to \infty$. We have therefore shown that V_1/V_0 is bounded by a constant times $\exp(-\frac{\pi^2 n}{T^2})$, which concludes the proof. $\qquad \square$

We now obtain an expression for the probability of staying in the band $[0, T]$, without touching the border, during a time $n \gg T^2$.

Lemma 4.3. *Assume that $T \to \infty$ and that $\lim_{n \to \infty} \frac{n}{T^2} = +\infty$. Then, we have:*

- *If T is odd,*

$$\mathbb{P}_z \left(\tau > n \right) = \frac{2}{T} \sin \left(\frac{z\pi}{T} \right) \frac{1}{\tan \left(\frac{\pi}{2T} \right)} e^{-ng(T)} \left(1 + \bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}}) \right). \tag{4.5}$$

- *If T is even, letting $a = \mathbb{1}_{\{n \leftrightarrow z\}}$*

$$\mathbb{P}_z \left(\tau > n \right) = \frac{4}{T} \sin \left(\frac{z\pi}{T} \right) \frac{\cos^a \left(\frac{\pi}{T} \right)}{\sin \left(\frac{\pi}{T} \right)} e^{-ng(T)} \left(1 + \bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}}) \right). \tag{4.6}$$

In particular, with a Taylor expansion, we get that

$$f_n(z, T) := \mathbb{P}_z \left(\tau > n \right) = \frac{4}{\pi} \sin \left(\frac{z\pi}{T} \right) e^{-ng(T)} \left[1 + \bar{\mathcal{O}}(T^{-2}) + \bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}}) \right], \tag{4.7}$$

and note that if $n \geq \frac{1}{4} T^2 \log T$ then $e^{-\frac{\pi^2 n}{T^2}} \leq T^{-\frac{1}{4}\pi^2} \leq T^{-2}$.

Proof. First of all, we write

$$\mathbb{P}_z \left(\tau > n \right) = \sum_{k > n} \left(\mathbb{P}_z \left(\tau = \tau_0 = k \right) + \mathbb{P}_z \left(\tau = \tau_T = k \right) \right). \tag{4.8}$$

When T is odd, then in (4.8), for each k in the sum there is only one term which is non-zero: applying Lemma 4.2 to estimate that term, we get

$$\mathbb{P}_z \left(\tau > n \right) = \left(1 + \bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}}) \right) \frac{2}{T} \sin \left(\frac{z\pi}{T} \right) \tan \left(\frac{\pi}{T} \right) \sum_{k > n} e^{-g(T)k}$$

$$= \left(1 + \bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}}) \right) \frac{2}{T} \sin \left(\frac{z\pi}{T} \right) \tan \left(\frac{\pi}{T} \right) e^{-g(T)n} \frac{\cos \left(\frac{\pi}{T} \right)}{1 - \cos \left(\frac{\pi}{T} \right)},$$

recalling that $e^{-g(T)} = \cos \left(\frac{\pi}{T} \right)$. This gives the desired result since $\frac{\sin(\theta)}{1 - \cos(\theta)} = \frac{1}{\tan(\theta/2)}$.

When T is even, notice that in (4.8), either $k \leftrightarrow z$ and then both terms are non-zero or $k \not\leftrightarrow z$ and then both terms are zero. Applying Lemma 4.2, we get

$$\mathbb{P}_z \left(\tau > n \right) = \left(1 + \bar{\mathcal{O}}(e^{-\frac{\pi^2 n}{T^2}}) \right) \frac{4}{T} \sin \left(\frac{z\pi}{T} \right) \tan \left(\frac{\pi}{T} \right) \sum_{k > n} e^{-g(T)k} \mathbb{1}_{\{k \leftrightarrow z\}}.$$

To deal with the last sum, denote $n^* = n^*(z) := \min\{k > n, k \leftrightarrow z\}$: note that n^* is equal to $n + 1 + a$ with $a = \mathbb{1}_{\{n \leftrightarrow z\}}$. The indices for which the term is not zero can be written as $k = n^* + 2j$ and thus

$$\sum_{k > n} e^{-g(T)k} \mathbb{1}_{\{k \leftrightarrow z\}} = e^{-n^* g(T)} \sum_{j \geq 0} e^{-2g(T)j} = e^{-ng(T)} \frac{\cos^{1+a} \left(\frac{\pi}{T} \right)}{1 - \cos^2 \left(\frac{\pi}{T} \right)},$$

recalling that $e^{-g(T)} = \cos \left(\frac{\pi}{T} \right)$. This gives the announced expression. $\qquad \square$

16

4.2 Range estimates

Recall the definition of the event $E_x^y(n) = \{M_n^- = -x, M_n^+ = y\}$ for any two non-negative integers x and y. We use Lemma 4.3 to estimate $\mathbb{P}_0(E_x^y(n))$, *i.e.* to prove Theorem 1.4. From this point onward, we always denote $T := x + y$. Using the spatial invariance of the random walk, we study the probability starting from x to stay in the strip $[0, T]$ *and* to touch both borders before time n. The symmetry of the walk allows us to assume $x \leq y$ and $0 \leq x \leq \frac{T}{2}$. We now write the probability of $E_x^y(n)$ as the following differences

$$
\begin{aligned}
\mathbb{P}_0\left(E_x^y(n)\right) &= \mathbb{P}_0\left(M_n^+ < y+1, M_n^- = -x\right) - \mathbb{P}_0\left(M_n^+ < y, M_n^- = -x\right) \\
&= \mathbb{P}_0\left(M_n^+ < y+1, M_n^- > -x-1\right) - \mathbb{P}_0\left(M_n^+ < y+1, M_n^- > -x\right) \\
&\quad - \left[\mathbb{P}_0\left(M_n^+ < y, M_n^- > -x-1\right) - \mathbb{P}_0\left(M_n^+ < y, M_n^- > -x\right)\right].
\end{aligned}
$$

Then each of those probabilities is of a strict confinement event with different strips widths and starting points: we get

$$
\mathbb{P}_0\left(E_x^y(n)\right) = f_n(x+1, T+2) - f_n(x, T+1) - f_n(x+1, T+1) + f_n(x, T). \tag{4.9}
$$

We can therefore use Lemma 4.3 to estimate each of these terms. We will find different asymptotics depending on whether $\frac{n}{T^3}$ goes to 0 or to $+\infty$ (or converges to a constant). This ratio is known to be the relevant quantity when studying such constrained random walk (see [8]). The main reason is that in the expansion (4.7) we have that $e^{-ng(T)} \sim e^{-ng(T+1)} \sim e^{-ng(T+2)}$ if and only if $\frac{n}{T^3}$ goes to 0 (see below): this case will prove to be more intricate because of several cancellations in (4.9).

4.2.1 First case: $\lim_{n \to \infty} \frac{n}{T^3} = +\infty$

In that case, recalling that $e^{-g(T)} = \cos(\frac{\pi}{T})$ we have

$$
\frac{e^{-g(T+1)n}}{e^{-g(T+2)n}} = \left(1 - (1 + \bar{o}(1))\frac{\pi^2}{T^3}\right)^n \xrightarrow{n \to \infty} 0, \quad \frac{e^{-g(T)n}}{e^{-g(T+2)n}} = \left(1 - (1 + \bar{o}(1))\frac{2\pi^2}{T^3}\right)^n \xrightarrow{n \to \infty} 0.
$$

Therefore, in view of (4.7), we have that $f_n(x, T+1), f_n(x+1, T+1), f_n(x, T)$ are all negligible compared to $f_n(x+1, T+2)$. Using (4.9), and (4.7), we therefore get that

$$
\mathbb{P}_0\left(E_x^y(n)\right) = (1 + \bar{o}(1))\frac{4}{\pi}\sin\left(\frac{\pi(x+1)}{T}\right)e^{-g(T+2)n}, \tag{4.10}
$$

where the $\bar{o}(1)$ depends only on T and is uniform in x. Using also that $\sin(\frac{\pi(x+1)}{T+2}) = (1 + \bar{\mathcal{O}}(\frac{1}{T}))\sin(\frac{\pi(x+1)}{T})$ uniformly in x, we get the desired result.

4.2.2 Second case: $\lim_{n \to \infty} \frac{n}{T^3} = \alpha \in (0, +\infty)$

Similarly as above, we have

$$
\lim_{n \to \infty} \frac{e^{-g(T+2)n}}{e^{-g(T)n}} = e^{2\alpha\pi^2}, \quad \lim_{n \to \infty} \frac{e^{-g(T+1)n}}{e^{-g(T)n}} = e^{\alpha\pi^2}.
$$

Therefore, using (4.9) and (4.7), we get that

$$
\frac{\pi}{4}e^{g(T)n}\mathbb{P}_0\left(E_x^y(n)\right)
$$
$$
= (1 + \bar{o}(1))\left[\sin\left(\frac{\pi(x+1)}{T}\right)e^{2\alpha\pi^2} + \sin\left(\frac{\pi x}{T}\right) - \left[\sin\left(\frac{\pi(x+1)}{T}\right) + \sin\left(\frac{\pi x}{T}\right)\right]e^{\alpha\pi^2}\right]
$$

where we have used that $\sin(\frac{\pi(x+1)}{T+2}) = (1+\bar{o}(1))\sin(\frac{\pi(x+1)}{T})$ with $\bar{o}(1)$ uniform in x, and similarly with $\sin(\frac{\pi(x+1)}{T+1})$, $\sin(\frac{\pi x}{T+1})$. This gives the announced asymptotics.

Let us stress that, in the case where $x = \bar{\mathcal{O}}(1)$, we get

$$\mathbb{P}_0\left(E_x^y(n)\right) = (1+\bar{o}(1))\frac{4}{T}(e^{\alpha\pi^2} - 1)(e^{\alpha\pi^2}(x+1) - x)e^{-g(T)n}. \qquad (4.11)$$

If on the other hand we have $x \to \infty$, then we have

$$\mathbb{P}_0\left(E_x^y(n)\right) = (1+\bar{o}(1))\frac{4}{\pi}(e^{\alpha\pi^2} - 1)^2 \sin\left(\frac{x\pi}{T}\right) e^{-g(T)n}, \qquad (4.12)$$

and one can also write $(e^{\alpha\pi^2} - 1)^2 = 4e^{\alpha\pi^2}\cosh^2(\alpha\pi^2)$.

4.2.3 Last case: $\lim_{n\to\infty}\frac{n}{T^3} = 0$

When $\lim_{n\to\infty}\frac{n}{T^3} = 0$, the first order terms all cancel each other and we have to look further in the asymptotic expansion. We will show that

$$\frac{\pi}{4}\mathbb{P}_0\left(E_x^y(n)\right)e^{g(T+1)n} = (1+\bar{o}(1))\frac{\pi^4 n^2}{T^6}\sin\left(\frac{\pi(x+\frac{1}{2})}{T}\right) + (1+\bar{o}(1))\frac{\pi^3 n}{T^4}\left(1 - \frac{2x}{T}\right)\cos\left(\frac{\pi x}{T}\right).$$

This proves the desired result since the second term can only be dominant if $x/T \to 0$; it is actually dominant in particular if $x = \bar{\mathcal{O}}(1)$.

Recalling that $\lim_{n\to\infty}\frac{n}{T^2} = +\infty$ (in particular $\frac{n}{T^4} = \bar{o}(\frac{n^2}{T^6})$), we have the following expansions:

$$\frac{e^{-g(T)n}}{e^{-g(T+1)n}} = 1 - \frac{\pi^2 n}{T^3} + (1+\bar{o}(1))\frac{\pi^4 n^2}{2T^6}, \qquad \frac{e^{-g(T+2)n}}{e^{-g(T+1)n}} = 1 + \frac{\pi^2 n}{T^3} + (1+\bar{o}(1))\frac{\pi^4 n^2}{2T^6}(1+\bar{o}(1)).$$

Hence, from (4.9), using (4.7) (with the fact that $n \geq \frac{1}{4}T^2\log T$ so that $\bar{\mathcal{O}}\left(e^{-\frac{\pi^2 n}{T^2}}\right) = \bar{o}(T^{-2})$) we get

$$\frac{\pi}{4}\mathbb{P}_0\left(E_x^y(n)\right)e^{g(T+1)n} = \left\{ \sin\left(\frac{\pi x}{T}\right)\left[1 - \frac{\pi^2 n}{T^3} + (1+\bar{o}(1))\frac{\pi^4 n^2}{2T^6}\right] \right.$$

$$- \left(\sin\left(\frac{\pi x}{T+1}\right) + \sin\left(\frac{\pi(x+1)}{T+1}\right)\right)\left[1 + \bar{\mathcal{O}}(T^{-2})\right]$$

$$\left. + \sin\left(\frac{\pi(x+1)}{T+2}\right)\left[1 + \frac{\pi^2 n}{T^3} + (1+\bar{o}(1))\frac{\pi^4 n^2}{2T^6}\right] \right\}.$$

Note that we absorbed all terms $\bar{\mathcal{O}}(T^{-2})$ in the $\bar{o}(\frac{n^2}{T^6})$, since $\lim_{n\to\infty}\frac{n}{T^2} = +\infty$. Hence, we get that

$$\frac{\pi}{4}\mathbb{P}_0\left(E_x^y(n)\right)e^{g(T+1)n} = (1+\bar{o}(1))\frac{\pi^4 n^2}{2T^6}\left(\sin\left(\frac{\pi x}{T}\right) + \sin\left(\frac{\pi(x+1)}{T}\right)\right) + A + B,$$

with

$$A = \sin\left(\frac{x\pi}{T}\right) - 2\sin\left(\frac{\pi(x+\frac{1}{2})}{T+1}\right) + \sin\left(\frac{\pi(x+1)}{T+2}\right),$$

$$B = \frac{2n\pi^2}{T^3}\left[\sin\left(\frac{\pi(x+1)}{T+2}\right) - \sin\left(\frac{\pi x}{T}\right)\right].$$

Here, for A, we have also used that $\sin(\frac{\pi x}{T+1}) + \sin(\frac{\pi(x+1)}{T+1}) = 2\sin(\frac{\pi(x+\frac{1}{2})}{T+1})\cos(\frac{\pi}{2(T+1)})$ with $\cos(\frac{\pi}{2(T+1)}) = 1 + \bar{\mathcal{O}}(T^{-2})$ and absorbed the $\bar{\mathcal{O}}(T^{-2})$ in the $\bar{o}(\frac{n^2}{T^6})$. We show below that A is negligible compared to B, so let us start by estimating B.

18

Term B.   Note that setting $v := \frac{T}{2} - x$ we have

$$B = \frac{2n\pi^2}{T^3}\left[\cos\left(\frac{\pi v}{T+2}\right) - \cos\left(\frac{\pi v}{T}\right)\right].$$

Using the formula for the difference of cosines, we get that

$$\cos\left(\frac{\pi v}{T+2}\right) - \cos\left(\frac{\pi v}{T+1}\right) = 2\sin\left(\frac{\pi v}{2(T+1)(T+2)}\right)\sin\left(\frac{\pi v}{T}\frac{T+\frac{3}{2}}{T+2}\right)$$

$$= (1+\bar{o}(1))\frac{\pi v}{T^2}\sin\left(\frac{\pi v}{T}\right).$$

We end up with

$$B = (1+\bar{o}(1))\frac{2n\pi^3 v}{T^5}\sin\left(\frac{\pi v}{T}\right).$$

Term A.   As far as A is concerned, notice that setting $v := \frac{T}{2} - x$ we have

$$A = \cos\left(\frac{\pi v}{T}\right) - 2\cos\left(\frac{\pi v}{T+1}\right) + \cos\left(\frac{\pi v}{T+2}\right).$$

Using the formula for the difference of cosines, we get that $A/2$ is equal to

$$-\sin\left(\frac{\pi v}{2T(T+1)}\right)\sin\left(\frac{\pi v}{T}\frac{T+\frac{1}{2}}{T+1}\right) + \sin\left(\frac{\pi v}{2(T+1)(T+2)}\right)\sin\left(\frac{\pi v}{T}\frac{T(T+\frac{3}{2})}{(T+1)(T+2)}\right)$$

$$= \frac{\pi v}{2T^2}\left[\sin\left(\frac{\pi v}{T}\frac{T(T+\frac{3}{2})}{(T+1)(T+2)}\right) - \sin\left(\frac{\pi v}{T}\frac{T+\frac{1}{2}}{T+1}\right)\right] + \bar{\mathcal{O}}\left(\frac{v}{T^3}\sin\left(\frac{\pi v}{T}\right)\right),$$

where we have used that $\sin(\frac{\pi v}{2T(T+1)}) = \frac{\pi v}{2T^2}(1+\bar{\mathcal{O}}(T^{-1}))$ and similarly for $\sin(\frac{\pi v}{2(T+1)(T+2)})$. Using the formula for the difference of sines, we get

$$\sin\left(\frac{\pi v}{T}\frac{T(T+\frac{3}{2})}{(T+1)(T+2)}\right) - \sin\left(\frac{\pi v}{T}\frac{T+\frac{1}{2}}{T+1}\right)$$

$$= -2\sin\left(\frac{\pi v}{2T(T+2)}\right)\cos\left(\frac{\pi v}{T}\left[1+\bar{\mathcal{O}}(T^{-1})\right]\right).$$

Hence, we end up with $A = \bar{\mathcal{O}}(\frac{v^2}{T^4}\cos(\frac{\pi v}{T})) + \bar{\mathcal{O}}(\frac{v}{T^3}\sin(\frac{\pi v}{T}))$. This is negligible compared to B, since $\lim_{n\to\infty}\frac{n}{T^2} = +\infty$.

4.2.4   Estimates for a positive random walk

Note that our estimates need adjustments when $x = 0$ or $x = T$. We go back to correct (4.9) in order to take into account that 0 is the starting point of the walk. We write

$$\mathbb{P}_0\left(E_0^y(n)\right) = \mathbb{P}_0\left(M_n^- > -1, M_n^+ < y+1\right) - \mathbb{P}_0\left(M_n^- > -1, M_n^+ < y\right)$$

$$= f_n(1, T+1) - f_n(1, T)$$

Note that $y = T$ but we will keep separating the notations y and T. thus, we have

$$\mathbb{P}_0\left(E_0^y(n)\right) = \frac{4}{\pi}\left[\sin\left(\frac{\pi}{T+2}\right)e^{-g(T+2)n} - \sin\left(\frac{\pi}{T+1}\right)e^{-g(T+1)n} + \bar{\mathcal{O}}(T^{-3})\right] \quad (4.13)$$

Once again we have different asymptotics depending on the ratio n/T^3 that we rapidly present in the following

Case $\frac{n}{T^3} \to +\infty$ As previously, $e^{-g(T+2)n}$ is the dominant term and thus

$$\mathbb{P}_0\left(E_0^y(n)\right) = \frac{4}{\pi}(1 + \bar{o}(1)) \sin\left(\frac{\pi}{T+2}\right) e^{-g(T+2)n} \tag{4.14}$$

And we get the formula of (4.10) applied to $x = 0$.

Case $\frac{n}{T^3} \to \alpha$ Factorize by $e^{-g(T)n}$ as in the general case, we thus write

$$\mathbb{P}_0\left(E_0^y(n)\right) = \frac{4}{\pi}\left[\sin\left(\frac{\pi}{T+2}\right) e^{2\alpha\pi^2} - \sin\left(\frac{\pi}{T+1}\right) e^{\alpha\pi^2} + \bar{\mathcal{O}}(T^{-3})\right] e^{-g(T)n}$$

That can we rewritten as

$$\mathbb{P}_0\left(E_0^y(n)\right) = \frac{4}{\pi}(1 + \bar{o}(1)) e^{\alpha\pi^2}(e^{\alpha\pi^2} - 1) \sin\left(\frac{\pi}{T}\right) e^{-g(T)n}$$

which is exactly (4.11) taken at $x = 0$.

Case $\frac{n}{T^3} \to 0$ In this case, we again factorize by $e^{-g(T+1)n}$ and write

$$\mathbb{P}_0\left(E_0^y(n)\right) = \frac{4}{\pi}\left[\sin\left(\frac{\pi}{T+2}\right)\left[1 + \frac{n\pi^2}{(T+1)^3} + \frac{n^2\pi^4}{2T^6}(1 + \bar{o}(1))\right]\right.$$
$$\left. - \sin\left(\frac{\pi}{T+1}\right) + \bar{\mathcal{O}}(T^{-3})\right] e^{-g(T+1)n}$$

We are left to compare all the terms in this expression :

$$A = \sin\left(\frac{\pi}{T+2}\right) - \sin\left(\frac{\pi}{T+1}\right) = -2\sin\left(\frac{\pi}{2(T+1)(T+2)}\right)\cos\left(\frac{\pi}{2T}\right) \sim -\frac{2\pi}{T^2}$$

$$B = \frac{n\pi^2}{(T+1)^3}\sin\left(\frac{\pi}{T+2}\right) \sim \frac{n\pi^3}{T^4}, \qquad D = \frac{n^2\pi^4}{T^6}\sin\left(\frac{\pi}{T+2}\right) \sim \frac{n^2\pi^5}{T^7}$$

See that $A \ll B$ and $D \ll B$ using both $\frac{n}{T^2} \to \infty$ and $\frac{n}{T^3} \to 0$, meaning that

$$\mathbb{P}_0\left(E_0^y(n)\right) = \frac{4n\pi}{(T+1)^3}(1 + \bar{o}(1)) \sin\left(\frac{\pi}{T+2}\right) e^{-g(T+1)n}$$

Acknowledgements

References

[1] N. Berestycki and R. Cerf. "The random walk penalised by its range in dimensions $d \geq 3$". In: *Annales Henri Lebesgue* 4 (2021), pp. 1–79. DOI: 10.5802/ahl.66. URL: https://ahl.centre-mersenne.org/articles/10.5802/ahl.66/.

[2] Q. Berger, F. Den Hollander, and J. Poisat. "Annealed scaling for a charged polymer in dimensions two and higher". In: *Journal of Physics A: Mathematical and Theoretical* 51.5 (2018). HAL-ID: hal-01576410.

[3]   Q. Berger, C.-H. Huang, N. Torri, and R. Wei. *One-dimensional polymers in random environments: stretching vs. folding*. 2021. arXiv: `2002.06899v3` [`math.PR`].

[4]   Q. Berger, N. Torri, and R. Wei. *Non-directed polymers in heavy-tail random environment in dimension $d \geq 2$*. 2021. arXiv: `2101.05949` [`math.PR`].

[5]   E. Bolthausen. "Localization of a two-dimensional random walk with an attractive path interaction". In: *The Annals of Probability* (1994), pp. 875–918.

[6]   N. Bouchot. *One dimensional polymer in a repulsive random environnement*. (Work in progress). 2022.

[7]   F. Caravenna, F. Den Hollander, N. Pétrélis, and J. Poisat. "Annealed scaling for a charged polymer". In: *Mathematical Physics, Analysis and Geometry* 19.1 (2016), pp. 1–87.

[8]   F. Caravenna and N. Pétrélis. "Depinning of a polymer in a multi-interface medium". In: *Electronic Journal of Probability* 14 (2009), pp. 2038–2067. DOI: `10.1214/EJP.v14-698`.

[9]   F. Den Hollander. *Random Polymers: École d'Été de Probabilités de Saint-Flour XXXVII–2007*. Springer, 2009.

[10]   J. Ding, R. Fukushima, R. Sun, and C. Xu. "Geometry of the random walk range conditioned on survival among Bernoulli obstacles". In: *Probability Theory and Related Fields* 177.1 (2020), pp. 91–145.

[11]   M. D. Donsker and S. S. Varadhan. "Asymptotics for the Wiener sausage". In: *Communications on Pure and Applied Mathematics* 28.4 (1975), pp. 525–565. DOI: `10.1002/cpa.3160280406`.

[12]   M. D. Donsker and S. S. Varadhan. "On the number of distinct sites visited by a random walk". In: *Communications on Pure and Applied Mathematics* 32.6 (1979), pp. 721–747. DOI: `10.1002/cpa.3160320602`.

[13]   W. Feller. *An Introduction to Probability Theory and its Applications*. 3rd ed. Vol. 1. John Wiley & Sons, 1968. Chap. XIV.5. ISBN: 978-0-471-25708-0.

[14]   C.-H. Huang. *The scaling limits for Wiener sausages in random environments*. 2019. arXiv: `1902.04930` [`math.PR`].

[15]   R. L. Schilling and L. Partzsch. *Brownian motion*. 2nd ed. de Gruyter, 2014. ISBN: 978-3110373981.

[16]   U. Schmock. "Convergence of the normalized one-dimensional wiener sausage path measures to a mixture of brownian taboo processes". In: *Stochastics and Stochastic Reports* 29.2 (1990), pp. 171–183. DOI: `10.1080/17442509008833612`.

LPSM, Sorbonne Université, UMR 8001, Campus Pierre et Marie Curie, Boîte courrier 158, 4 Place Jussieu, 75252 Paris Cedex 05, France.

E-mail address: `nicolas.bouchot@sorbonne-universite.fr`