The Annals of Probability
2016, Vol. 44, No. 5, 3234–3290
DOI: 10.1214/15-AOP1046

INTERACTING PARTIALLY DIRECTED SELF AVOIDING WALK. FROM PHASE TRANSITION TO THE GEOMETRY OF THE COLLAPSED PHASE

BY PHILIPPE CARMONA*, GIA BAO NGUYEN* AND NICOLAS PÉTRÉLIS†

Université de Nantes and Universidad de Chile*†

In this paper, we investigate a model for a $1+1$ dimensional self-interacting and partially directed self-avoiding walk, usually referred to by the acronym IPDSAW. The interaction intensity and the free energy of the system are denoted by β and f, respectively. The IPDSAW is known to undergo a collapse transition at β_c. We provide the precise asymptotic of the free energy close to criticality, that is, we show that $f(\beta_c - \varepsilon) \sim \gamma \varepsilon^{3/2}$ where γ is computed explicitly and interpreted in terms of an associated continuous model. We also establish some path properties of the random walk inside the collapsed phase ($\beta > \beta_c$). We prove that the geometric conformation adopted by the polymer is made of a succession of long vertical stretches that attract each other to form a unique macroscopic bead and we establish the convergence of the region occupied by the path properly rescaled toward a deterministic Wulff shape.

1. Introduction.

1.1. *Model and physical insight.* A solvent is said to be "poor" for a given homopolymer if the chemical affinity between the solvent and the monomers constituting the homopolymer is low. When dipped in such a solvent, the homopolymer folds itself up to exclude the solvent and, therefore, adopts a collapsed conformation that looks like a compact ball. If the quality of the solvent improves, the chemical affinity raises until it reaches a threshold above which the polymer extends itself in such a way that a positive fraction of its monomers are in contact with the solvent.

The interacting partially directed self-avoiding walk (IPDSAW) was introduced in Zwanzig and Lauritzen (1968) as a partially directed model of an homopolymer in a poor solvent. The spatial configurations of the polymer of length L (L monomers) are modeled by the trajectories of a *self-avoiding* random walk on $\mathbb{Z}^2$ that only takes unitary steps *upward, downward and to the right*. Thus, the set of

Received February 2014; revised July 2015.
MSC2010 subject classifications. Primary 60K35; secondary 82B26, 82B41.
Key words and phrases. Polymer collapse, phase transition, variational formula, Wulff shape.

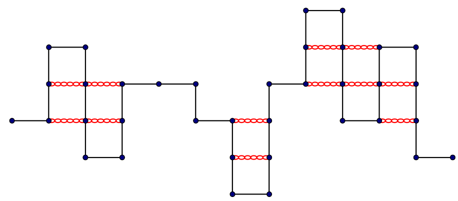

FIG. 1. *Example of a trajectory with* 12 *self-touchings in light grey.*

allowed L-step paths is

$$\mathcal{W}_L = \{w = (w_i)_{i=0}^L \in (\mathbb{N}_0 \times \mathbb{Z})^{L+1} : w_0 = 0, w_L - w_{L-1} = \rightarrow,\\ w_{i+1} - w_i \in \{\uparrow, \downarrow, \rightarrow\} \ \forall 0 \le i < L-1,\\ w_i \ne w_j \ \forall i < j\}.$$

Note that the choice of w ending with an horizontal step is made for convenience only. We consider two different a priori laws on $\mathcal{W}_L$, uniform and nonuniform.

(1) The uniform model: all L-step paths have the same probability, that is,

$$\mathbf{P}_L(w) = \frac{1}{|\mathcal{W}_L|}, \qquad w \in \mathcal{W}_L. \tag{1.1}$$

(2) The nonuniform model: the L-step paths have the following law:

- At the origin or after an horizontal step: the walker must step north, south or east with equal probability $1/3$.
- After a vertical step north (resp. south): the walker must step north (resp. south) or east with probability $1/2$.

Henceforth, we will focus on the uniform model since all our results can be adapted straightforwardly to the nonuniform model modulo a shift in the critical point β_c and in the value of the constant a_β defined before the shape theorem.

The monomer-solvent interactions are not taken into account directly in the IPDSAW. We rather consider that, when dipped in a poor solvent, the monomers try to exclude the solvent and, therefore, attract one another. For this reason, any nonconsecutive vertices of the walk though adjacent on the lattice are called *self-touchings* (see Figure 1) and the interactions between monomers are taken into account by assigning an energetic reward $\beta \ge 0$ to the polymer for each self-touching (consequently, a lower chemical affinity corresponds to a larger β). Thus, we associate with every random walk trajectory $w = (w_i)_{i=0}^L \in \mathcal{W}_L$ the Hamiltonian

$$H_L(w) := \sum_{\substack{i,j=0\\ i<j-1}}^{L} \mathbf{1}_{\{\|w_i - w_j\| = 1\}}, \tag{1.2}$$

which allows to define the law $P_{L,\beta}$ of the polymer in size L as

$$P_{L,\beta}(w) = \frac{e^{\beta H_{L,\beta}(w)}}{Z_{L,\beta}} \mathbf{P}_L(w), \tag{1.3}$$

where $Z_{L,\beta}$ is the normalizing constant known as the partition function of the system. Henceforth, we remove the term $1/|\mathcal{W}_L|$ from the definition of $\mathbf{P}_L$ [recall (1.1)] and from the computation of the partition function $Z_{L,\beta}$. Although $\mathbf{P}_L$ is not a probability law anymore, the latter simplification is harmless, because it does not change the polymer law $P_{L,\beta}$ and because it only induces a constant shift of the free energy $f(\beta)$ introduced in Section 1.2 below.

1.1.1. *From random walk paths to vertical stretches.* It is easy to see that any path in $\mathcal{W}_L$ can be decomposed into a collection of vertical stretches separated by one horizontal step. Thus, we set $\Omega_L := \bigcup_{N=1}^{L} \mathcal{L}_{N,L}$, where $\mathcal{L}_{N,L}$ is the set of all possible configurations consisting of N vertical stretches that have a total length L, that is,

$$\mathcal{L}_{N,L} = \left\{ l \in \mathbb{Z}^N : \sum_{n=1}^{N} |l_n| + N = L \right\}. \tag{1.4}$$

We build the natural one to one correspondence between Ω_L and $\mathcal{W}_L$ by associating with a given $l \in \Omega_L$ the path of $\mathcal{W}_L$ that starts at 0, takes $|l_1|$ vertical steps north if $l_1 > 0$ and south if $l_1 < 0$, then takes one horizontal step, then takes $|l_2|$ vertical steps north if $l_2 > 0$ and south if $l_2 < 0$ then takes one horizontal step and so on... (see Figure 2). The Hamiltonian associated with a given path of $\mathcal{W}_L$ can be rewritten in terms of its associated collection of vertical stretches $l \in \Omega_L$ as

$$H_L(l_1, \ldots, l_N) = \sum_{n=1}^{N-1} (l_n \,\tilde{\wedge}\, l_{n+1}), \tag{1.5}$$

where

$$x \,\tilde{\wedge}\, y = \begin{cases} |x| \wedge |y|, & \text{if } xy < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{1.6}$$

Therefore, the partition function can be rewritten under the form

$$Z_{L,\beta} = \sum_{N=1}^{L} \sum_{l \in \mathcal{L}_{N,L}} e^{\beta \sum_{i=1}^{N-1} (l_i \tilde{\wedge} l_{i+1})}. \tag{1.7}$$

1.2. *Free energy and collapse transition.* The sequence $\{\log Z_{L,\beta}\}_L$ is super-additive and the Hamiltonian in (1.2) is obviously bounded from above by βL. As a consequence, we can define the free energy per step $f : (0, \infty) \to \mathbb{R}$ as

$$f(\beta) = \lim_{L \to \infty} \frac{1}{L} \log Z_{L,\beta} = \sup_{L \in \mathbb{N}} \frac{1}{L} \log Z_{L,\beta} \le \beta. \tag{1.8}$$

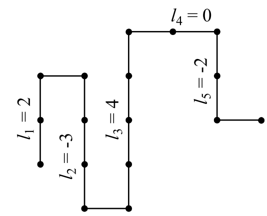

FIG. 2. *Example of a trajectory with $N = 5$ vertical stretches and length $L = 16$.*

The collapse transition corresponds to a loss of analyticity of $\beta \mapsto f(\beta)$ at some critical parameter $\beta_c \in (0, \infty)$ above which the density of self-touchings performed by the polymer equals 1. In this collapsed phase, the expression of the free energy per step is rather simple, that is, $\beta + \kappa$, where κ is the entropic constant associated to those trajectories in $\mathcal{W}_L$ whose self-touching density is equal to $1 + o(1)$. To achieve such a saturation of its self-touching, the polymer must choose its configuration among those satisfying two major geometric restrictions, that is,

- the number of horizontal steps is $o(L)$;
- most pairs of consecutive vertical stretches are of opposite directions.

It turns out that an appropriate choice of a trajectory satisfying both restrictions above is sufficient to exhibit the collapsed free energy. To that aim, we pick $L \in \mathbb{N}$: $\sqrt{L} \in \mathbb{N}$ and consider the trajectory $l^* \in \mathcal{L}_{\sqrt{L},L}$ defined as $l_i^* = (-1)^{i-1}(\sqrt{L} - 1)$ for $i \in \{1, \dots, \sqrt{L}\}$. By computing the contribution of l^* to $Z_{L,\beta}$ one immediately obtain that,[1] for $\beta > 0$,

$$f(\beta) \geq \beta. \tag{1.9}$$

At this stage, we can define the *excess free energy* $\tilde{f}(\beta) := f(\beta) - \beta$, which is always nonnegative by (1.9). We define the critical parameter

$$\beta_c := \inf\{\beta \geq 0 : \tilde{f}(\beta) = 0\}, \tag{1.10}$$

and the convexity of $\beta \mapsto \tilde{f}(\beta)$ allows us to partition $[0, \infty)$ into a collapsed phase denoted by $\mathcal{C}$ and an extended phase denoted by $\mathcal{E}$, that is,

$$\mathcal{C} := \{\beta : \tilde{f}(\beta) = 0\} = \{\beta : \beta \geq \beta_c\} \tag{1.11}$$

and

$$\mathcal{E} := \{\beta : \tilde{f}(\beta) > 0\} = \{\beta : \beta < \beta_c\}. \tag{1.12}$$

[1] In a previous paper, Nguyen and Pétrélis (2013) the authors obtained the lower bound of $f(\beta) \geq \beta - \log(1 + \sqrt{2})$. The difference comes from the omission of the normalizing factor $1/|\mathcal{W}_L|$.

1.3. *Main results*. The main results of this paper are Theorems A, B, C, D, E and F. Theorems A and B are dedicated to the investigation of the phase transition while the path properties of the polymer inside its collapsed phase are studied with Theorems C, D, E and F.

Before stating the theorems, we need to introduce $\mathbf{P}_\beta$ the law of an auxiliary symmetric random walk $V := (V_n)_{n\in\mathbb{N}}$ with geometric increments, that is, $V_0 = 0$, $V_n = \sum_{i=1}^n U_i$ for $n \in \mathbb{N}$ and $(U_i)_{i\in\mathbb{N}}$ is an i.i.d. sequence under the law $\mathbf{P}_\beta$, with distribution

$$\mathbf{P}_\beta(U_1 = k) = \frac{e^{-(\beta/2)|k|}}{c_\beta} \qquad \forall k \in \mathbb{Z} \text{ with } c_\beta := \frac{1+e^{-\beta/2}}{1-e^{-\beta/2}}. \tag{1.13}$$

Then, for $\delta \geq 0$ we set

$$\mathfrak{h}_\beta(\delta) := \lim_{N\to\infty} \frac{1}{N} \log \mathbf{E}_\beta\big(e^{-\delta A_N(V)}\big), \tag{1.14}$$

where $A_N(V) := \sum_{i=1}^N |V_i|$ gives the geometric area below the V trajectory after N steps. We will prove in Section 2.2 below that the limit in (1.14) exists and that $\delta \mapsto \mathfrak{h}_\beta(\delta)$ is nonpositive, nonincreasing and continuous on $[0,\infty)$. We finally define $\Gamma(\beta)$ an energetic term of crucial importance as

$$\Gamma(\beta) = \frac{c_\beta}{e^\beta}, \tag{1.15}$$

and we will see, for instance, in (1.36) below that $\Gamma(\beta)$ penalizes the horizontal steps when it is smaller than 1 and favors them when it is larger than 1.

1.3.1. *A sharper asymptotic of the free energy close to criticality.* With Theorem A, we give a new expression of the excess free energy.

THEOREM A (Free energy equation). *The excess free energy $\widetilde{f}(\beta)$ is the unique solution of the equation* $\log(\Gamma(\beta)) - \delta + \mathfrak{h}_\beta(\delta) = 0$ *if such a solution exists and* $\widetilde{f}(\beta) = 0$ *otherwise.*

Note that Theorem A and the obvious equality $\mathfrak{h}_\beta(0) = 0$ are sufficient to check that the critical parameter β_c is the unique solution of $\Gamma(\beta) = 1$. One of the main interest of Theorem A is that it allows us to use the analytic properties of $\delta \mapsto \mathfrak{h}_\beta(\delta)$ at 0^+ to investigate the regularity of $\beta \mapsto \widetilde{f}(\beta)$ at β_c.

THEOREM B (Phase transition asymptotics). *The phase transition is second order with critical exponent* $3/2$ *and the first order asymptotic of the excess free energy at* $(\beta_c)^-$ *is given by*

$$\lim_{\varepsilon\to 0^+} \frac{\widetilde{f}(\beta_c - \varepsilon)}{\varepsilon^{3/2}} = \left(\frac{\varsigma_1}{\varsigma_2}\right)^{3/2}, \tag{1.16}$$

where

$$\varsigma_1 = 1 + \frac{e^{-\beta_c/2}}{1 - e^{-\beta_c}}, \tag{1.17}$$

and where

$$\varsigma_2 = -\lim_{T\to\infty} \frac{1}{T} \log \mathbf{E}\big(e^{-\sigma_{\beta_c} \int_0^T |B(t)|\, dt}\big) = 2^{-1/3} |a_1'| \sigma_{\beta_c}^{2/3}, \tag{1.18}$$

with $\sigma_\beta^2 = \mathbf{E}_\beta(U_1^2)$, *with* a_1' *the smallest zero (in absolute value) of the first derivative of the Airy function and with* $(B_s)_{s\in[0,\infty)}$ *a standard Brownian motion.*

REMARK 1.1. The Laplace transform $\mathbf{E}(e^{-s\int_0^1 |B_s|\, ds})$ for $s > 0$ was first computed analytically in Kac (1946) and studied by Takács (1993) [see, e.g., the survey by Janson (2007)].

REMARK 1.2. The critical exponent $3/2$ is given by the leading term of the Taylor expansion of $\mathfrak{h}_\beta$ at 0^+, that is, $\mathfrak{h}_\beta(\gamma) \sim -c\gamma^{2/3}$ (with $c > 0$). The method of proof we used consists in cutting the trajectories into blocs of size $\gamma^{-2/3}$. This very method was used in van der Hofstad, den Hollander and König (2003), in dimension $d = 1$, to prove that discrete Domb–Joyce- type models converge toward continuous Edwards-type models in the weak coupling limit.

REMARK 1.3. The asymptotic $\mathfrak{h}_\beta(\gamma) \sim -c\gamma^{2/3}$ is closely related to the investigation of the so-called pre-wetting phenomenon [see Hryniv and Velenik (2004), where the scaling exponent is obtained from a renormalization procedure similar to ours]. The pre-wetting phenomenon is observed when a thermodynamically stable gas is in contact with a substrate (hard-wall) that has a strong preference for the liquid phase. In such a situation, a thin layer of liquid may appear that separates the substrate from the gas. When the temperature T gets closer to the liquid/gas boiling temperature T_b, the layer of liquid becomes thicker. The liquid-gas interface can therefore be modeled by a random walk trajectory constrained to remain positive and whose area is penalized via a Gibbs factor $\delta A_N(V)$ where δ vanishes as $T \to T_b$. Close to criticality ($\delta = 0$), the correlation length of the system varies as $\delta^{-2/3}$ which explains the $2/3$ exponent of $\mathfrak{h}_\beta$ at 0^+.

The determination of the precise asymptotics of the free energy close to β_c brings the IPDSAW into a thin class of statistical mechanical models for which the behavior of the free energy close to criticality is well understood. This is the case, for instance, for the pinning/wetting model [see Giacomin (2011), Chapter 2]. Perturbing such models by adding a weak random component to their interactions is physically relevant [see Derrida, Hakim and Vannimenus (1992)] and gives rise to complex mathematical issues [see Alexander and Sidoravicius (2006)]. For the model of a polymer pinned by a linear interface, the issue of the disorder relevance

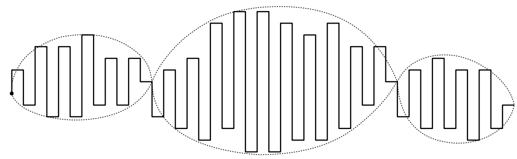

FIG. 3. *Example of a trajectory with* 3 *beads.*

on the phase transition was controversial until it was settled recently [see Derrida et al. (2009) or Giacomin (2011), Chapters 4 and 5, for a survey]. For the IPD-SAW, a natural way of introducing the disorder would be to assign an energetic price $\beta + s\xi_{i,j}$ to the self-touching between monomers i and j. The mechanism governing the phase transition being quite different from its counterpart in the pinning model, the investigation of the disorder effect is relevant both mathematically and physically.

1.3.2. *Path properties inside the collapsed phase.* The main result of this paper is concerned with the path behavior of the polymer inside its collapsed phase ($\beta > \beta_c$). We divide each trajectory into a succession of beads. Each bead is made of vertical stretches of strictly positive length and arranged in such a way that two consecutive stretches have opposite directions (north and south) and are separated by one horizontal step (see Figure 3). A bead ends when the polymer gives the same direction to two consecutive vertical stretches or when a zero length stretch appears, which corresponds to two consecutive horizontal steps. We will prove that the polymer folds itself up into a *unique macroscopic bead* and we will identify its horizontal extension and its asymptotic deterministic shape. To quantify these results, we need the following notation.

1.3.3. *Number of beads.* Let $l \in \Omega_L$ and denote by $N_L(l)$ its horizontal extension, that is, $N_L(l)$ is the integer N such that $l \in \mathcal{L}_{N,L}$. Pick $l \in \mathcal{L}_{N,L}$ and let $(u_j)_{j=1}^N$ be the sequence of accumulated lengths of the polymer after each vertical stretch, adding the lengths of the one step horizontal steps, that is $u_j = |l_1| + \cdots + |l_j| + j$ for $j \in \{1, \dots, N\}$. For convenience only, set $l_{N+1} = 0$. Set also $x_0 = 0$ and for $j \in \mathbb{N}$ such that $x_{j-1} < N$, set $x_j = \inf\{i \ge x_{j-1} + 1 : l_i \,\widetilde{\curlywedge}\, l_{i+1} = 0\}$ (see Figure 4). Finally, let $n_L(l)$ be the index of the last x_j that is well defined, that is, $x_{n_L(l)} = N$. Thus, we can decompose any trajectory $l \in \Omega_L$ into a succession of $n_L(l)$ beads, each of them being associated with a subinterval of $\{1, \dots, L\}$ written as

$$I_j = \{u_{x_{j-1}} + 1, \dots, u_{x_j}\} \qquad \text{for } j \in \{1, \dots, n_L(l)\} \tag{1.19}$$

and, therefore, we can partition $\{1, \dots, L\}$ into $\bigcup_{j=1}^{n_L(l)} I_j$. At this stage, we can define the largest bead of a trajectory $l \in \Omega_L$ as $I_{j_{\max}}$ with

$$j_{\max} = \arg\max\{|I_j|, j \in \{1, \dots, n_L(l)\}\}. \tag{1.20}$$

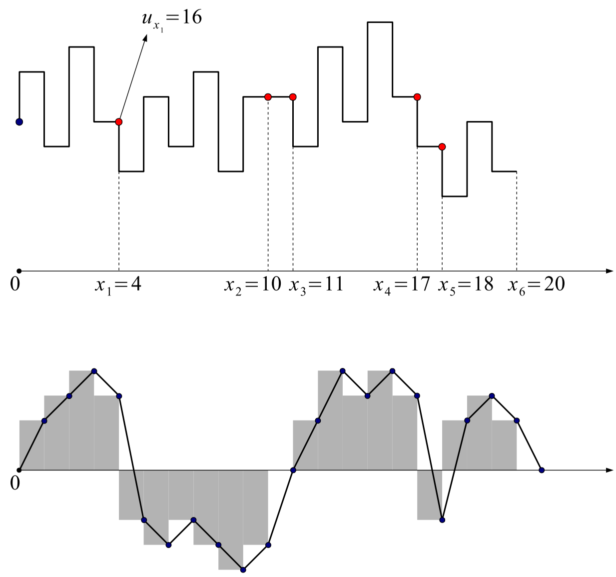

FIG. 4. *An example of a trajectory $l = (l_i)_{i=1}^{20}$ with 6 beads is drawn on the upper picture. The auxiliary random walk V associated with l, that is, $(V_i)_{i=0}^{21} = (T_{20})^{-1}(l)$ is drawn on the lower picture.*

With Theorem C below, we claim that, in the collapsed phase, there is only one macroscopic bead.

THEOREM C (One bead theorem). *For $\beta > \beta_c$, there exists a $c > 0$ such that*

$$\lim_{L\to\infty} P_{L,\beta}\big(|I_{j_{\max}}| \geq L - c(\log L)^4\big) = 1. \tag{1.21}$$

REMARK 1.4. Dividing trajectories into beads does not give rise to an underlying renewal process as, for instance, for the homogeneous pinning model when the trajectory is divided into excursions away from the origin [see, e.g., Giacomin (2007), Chapter 2]. The fact that, after a bead of length 1 the first stretch of the following bead can be either positive or negative whereas its orientation is constrained when the former bead is strictly larger than 1 creates a dependency between consecutive beads that prevents us from rewriting the partition function with the help of an associated renewal process. However, if we omit the dependency between consecutive beads then, thanks to Proposition 4.2, the "bead process" $(u_{x_j})_{j=0}^{n_L(l)}$ under $P_{L,\beta}$ can be related to a sub-exponential defective renewal process $\tau = (\tau_i)_{i\geq 0}$ conditioned on $L \in \tau$. This latter process is characterized by an inter-arrival law $K : \overline{\mathbb{N}} \to [0, 1]$ that satisfies $K(\infty) > 0$ and $K(n) = k(n)e^{-c\sqrt{n}}$ with $k : \mathbb{N} \to \mathbb{N}$ a slowly varying function. Once conditioned by $\{L \in \tau\}$, it can be proven [see Giacomin (2007), Appendix A.5 for the heavy tailed case or more recently Torri

(2014) where the sub-exponential case is explicitly treated] that the number of renewals is $O(1)$ and that again there is only one macroscopic renewal [see, e.g., Asmussen (2003) for a general background on renewal theory].

1.3.4. *Shape theorem.* First, recall the one-to-one correspondence between Ω_L and $\mathcal{W}_L$ described in Section 1.1 and denote by w_l the path in $\mathcal{W}_L$ associated with a given family of vertical stretches $l \in \Omega_L$. Then, identify each $l \in \Omega_L$ with a connected compact subset of $\mathbb{R}^2$ denoted by $S_L(l)$ that extends the sites of $\mathbb{Z}^2$ occupied by w_l to squares of length 1, that is,

$$S_L(l) = \left\{ \bigcup_{i=0}^{L} w_l(i) + \left[-\frac{1}{2}, \frac{1}{2}\right]^2 \right\}, \qquad l \in \Omega_L. \tag{1.22}$$

With Theorem D below, we prove that, once rescaled horizontally and vertically by $\sqrt{L}$ the subset $S_L(l)$ converges in probability and for the Hausdorff distance toward $\mathcal{S}_\beta$ a deterministic subset of $\mathbb{R}^2$. Before defining $\mathcal{S}_\beta$, we need to settle some notation.

First, we denote by $\mathfrak{L}(h)$, $h \in (-\frac{\beta}{2}, \frac{\beta}{2})$ the logarithmic moment generating function of the random variable U_1, that is,

$$\mathfrak{L}(h) := \log \mathbf{E}_\beta[e^{hU_1}], \tag{1.23}$$

and we introduce $\mathfrak{L}_\Lambda$

$$\mathfrak{L}_\Lambda(\mathbf{h}) := \int_0^1 \mathfrak{L}(xh_0 + h_1)\, dx, \tag{1.24}$$

which is defined on

$$\mathcal{D} := \left\{ \mathbf{h} = (h_0, h_1) \in \mathbb{R}^2 : h_1 \in \left(-\frac{\beta}{2}, \frac{\beta}{2}\right), h_0 + h_1 \in \left(-\frac{\beta}{2}, \frac{\beta}{2}\right) \right\}. \tag{1.25}$$

Then we let $\widetilde{\mathbf{h}}(q, 0) := (\widetilde{h}_0(q, 0), \widetilde{h}_1(q, 0))$ be the unique solution of the equation

$$\nabla \mathfrak{L}_\Lambda(\mathbf{h}) = (q, 0). \tag{1.26}$$

Since for $\beta > \beta_c$, the function

$$\widetilde{G}(a) := a \log \Gamma(\beta) - \frac{1}{a} \widetilde{h}_0\left(\frac{1}{a^2}, 0\right) + a \mathfrak{L}_\Lambda\left(\widetilde{\mathbf{h}}\left(\frac{1}{a^2}, 0\right)\right), \tag{1.27}$$

defined on $(0, \infty)$ is C^∞, strictly concave and negative (see Section 4.4), we let $a_\beta > 0$ be its unique maximizer.

We let γ_β^* be the Wulff shape minimizing the rate function of Mogulskii large deviation principle [see Dembo and Zeitouni (2010), Theorem 5.1.2] applied to the random walk of law $\mathbf{P}_\beta$, on the set containing the cadlag functions $\gamma : [0, 1] \to \mathbb{R}$

satisfying $\gamma(1) = 0$ and $\int_0^1 \gamma(t)\,dt = 1/a_\beta^2$ and endowed with the supremum norm $\|\cdot\|_\infty$. The following explicit formula holds (see Section 4.5):

$$\gamma_\beta^*(s) = \int_0^s \mathfrak{L}'\left[\left(\frac{1}{2} - x\right)\widetilde{h}_0\left(\frac{1}{a_\beta^2}, 0\right)\right] dx, \qquad s \in [0, 1]. \tag{1.28}$$

Eventually, we define the limiting shape

$$\mathcal{S}_\beta = \left\{(x, y) \in \mathbb{R}^2 : x \in [0, a_\beta], y \in \left[-\tfrac{1}{2} a_\beta \gamma_\beta^*(x/a_\beta), \tfrac{1}{2} a_\beta \gamma_\beta^*(x/a_\beta)\right]\right\} \tag{1.29}$$

and we denote by d_H the Hausdorff distance between subsets of $\mathbb{R}^2$.

THEOREM D (Shape theorem). *For $\beta > \beta_c$, we have convergence in $P_{L,\beta}$ probability for the Hausdorff distance toward a deterministic shape*

$$\lim_{L \to \infty} P_{L,\beta}\left(d_H\left(\frac{S_L(l)}{\sqrt{L}}, \mathcal{S}_\beta\right) > \varepsilon\right) = 0 \qquad (\forall \varepsilon > 0). \tag{1.30}$$

This shape theorem is equivalent to the combination of Theorems E and F below. We display in Appendix A a proof of this equivalence.

THEOREM E (Horizontal extension). *For $\beta > \beta_c$, for all $\varepsilon > 0$*

$$\lim_{L \to \infty} P_{L,\beta}\left(\left|\frac{N_L(l)}{\sqrt{L}} - a_\beta\right| > \varepsilon\right) = 0. \tag{1.31}$$

REMARK 1.5. Determining the horizontal extension is challenging in the extended regime ($\beta < \beta_c$) and in the critical regime ($\beta = \beta_c$) as well. In the extended regime, we can already derive from the variational formula of the free energy in Nguyen and Pétrélis (2013), Theorem 1.2, that there exists $c_2 > c_1 > 0$ so that $\lim_{L \to \infty} P_{L,\beta}(N_L(l)/L \in [c_1, c_2]) = 1$. The extension is therefore of order L and we expect that a law of large numbers also holds so that $N_L(l)/L$ converges in $P_{L,\beta}$ probability toward some constant $e_\beta \in (0, 1)$. The critical regime is more delicate. In view of the random walk representation and since $\Gamma(\beta_c) = 1$, the law of $N_L(l)$ when l is sampled from $P_{L,\beta}$ is exactly that of the stopping time $\tau_L := \inf\{n \geq 1 : n + A_n(V) \geq L\}$ of a random walk V of law $\mathbf{P}_\beta$ conditioned on $\{V_{\tau_L} = 0, A_{\tau_L} = L - \tau_L\}$. We expect that a Donsker-type invariance principle will hold there so that typically $A_{\tau_L} \sim \tau_L^{3/2}$ and thus we expect $N_L(l)/L^{2/3}$ to be tight under $P_{L,\beta}$.

The next theorem gives the scaling limit of the upper and lower envelopes of the path in the collapsed phase. Pick $l \in \mathcal{L}_{N,L}$ and let $\mathcal{E}_l^+ = (\mathcal{E}_{l,i}^+)_{i=0}^{N+1}$ be the path that links the top of each stretch consecutively (see Figure 5), while $\mathcal{E}_l^- = (\mathcal{E}_{l,i}^-)_{i=0}^{N+1}$

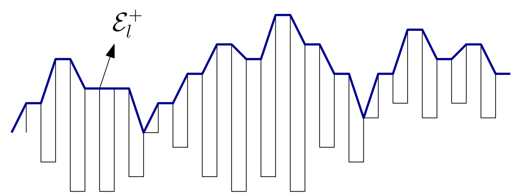

FIG. 5. *Example of the upper envelope of a trajectory.*

is the counterpart of $\mathcal{E}_l^+$ that links the bottom of each stretch consecutively. Thus, $\mathcal{E}_{l,0}^+ = \mathcal{E}_{l,0}^- = 0$,

$$\mathcal{E}_{l,i}^+ = \max\{l_1 + \cdots + l_{i-1}, l_1 + \cdots + l_i\}, \qquad i \in \{1, \dots, N\}, \tag{1.32}$$

$$\mathcal{E}_{l,i}^- = \min\{l_1 + \cdots + l_{i-1}, l_1 + \cdots + l_i\}, \qquad i \in \{1, \dots, N\}, \tag{1.33}$$

and $\mathcal{E}_{l,N+1}^+ = \mathcal{E}_{l,N+1}^- = l_1 + \cdots + l_N$. Then let $\widetilde{\mathcal{E}}_l^+$ and $\widetilde{\mathcal{E}}_l^-$ be the time–space rescaled cadlag processes associated with $\mathcal{E}_l^+$ and $\mathcal{E}_l^-$ and defined as

$$\begin{aligned} \widetilde{\mathcal{E}}_l^+(t) &= \frac{1}{N+1}\mathcal{E}_{l,\lfloor t(N+1)\rfloor}^+ \quad \text{and} \\ \widetilde{\mathcal{E}}_l^-(t) &= \frac{1}{N+1}\mathcal{E}_{l,\lfloor t(N+1)\rfloor}^-, \qquad t \in [0,1]. \end{aligned} \tag{1.34}$$

THEOREM F (Wulff shape). *For $\beta > \beta_c$ and $\varepsilon > 0$,*

$$\begin{aligned} \lim_{L\to\infty} P_{L,\beta}\left(\left\|\widetilde{\mathcal{E}}_l^+ - \frac{\gamma_\beta^*}{2}\right\|_\infty > \varepsilon\right) &= 0, \\ \lim_{L\to\infty} P_{L,\beta}\left(\left\|\widetilde{\mathcal{E}}_l^- + \frac{\gamma_\beta^*}{2}\right\|_\infty > \varepsilon\right) &= 0. \end{aligned} \tag{1.35}$$

Note that $\widetilde{\mathcal{E}}_l^+ - \widetilde{\mathcal{E}}_l^-$ [resp., $(\widetilde{\mathcal{E}}_l^+ + \widetilde{\mathcal{E}}_l^-)/2$] is the rescaled version of the process that associates with each index $i \in \{1, \dots, N_L(l)\}$ the length $|l_i|$ of the ith stretch (resp., the height of the middle of the ith stretch $l_1 + \cdots + l_{i-1} + \frac{l_i}{2}$). In view of Theorem F, the Wulff shape γ_β^* happens to be the limit, as $L \to \infty$, of $\widetilde{\mathcal{E}}_l^+ - \widetilde{\mathcal{E}}_l^-$. Such Wulff shape was identified, for instance, in Dobrushin and Hryniv (1996), as the limit of a random walk trajectory conditioned by fixing a large algebraic area between the path and the x-axis. However, the latter convergence is not sufficient to prove (1.35). We must indeed show that $(\widetilde{\mathcal{E}}_l^+ + \widetilde{\mathcal{E}}_l^-)/2$ converges to 0 in probability.

REMARK 1.6. The Wulff shape construction, initially displayed in Wulff (1901) appears in many models of statistical mechanics to describe the limiting shape of properly rescaled interfaces separating pure phases. Their construction is achieved by minimizing the integral of the surface tension along the continuous contours that satisfy some particular geometric constraint. A famous example arises from 2D Ising model in the phase transition regime. When considering a large square box of size N with boundary condition and $T < T_c$, and by conditioning the total magnetization to be shifted from its mean $(-m^*N^2)$ by a factor $a_N \gg N^{4/3}$, it was proven in Dobrushin, Kotecký and Shlosman (1992) at low temperature and then in Ioffe (1994, 1995) and Ioffe and Schonmann (1998) up to T_c that this magnetization shift is due to a unique $+$ island whose boundary, once rescaled by $1/\sqrt{a_N}$, converges toward a Wulff shape.

1.4. *Relationship to earlier work.* The IPDSAW and its continuous versions have attracted a lot of attention from *physicists* until very recently [see, e.g., Brak et al. (2009) or Samanta and Thirumalai (2013)]. The main method that has been employed to investigate the IPDSAW involves combinatorial techniques [see Brak, Guttmann and Whittington (1992), Brak, Owczarek and Prellberg (1993) or more recently Owczarek and Prellberg (2007)]. To be more specific, this method consists in providing an analytic expression of the generating function $G(z) = \sum_{L=1}^{\infty} Z_{L,\beta} z^L$ whose radius of convergence R satisfies $f = -\log R$. For a detailed version of the computations, we refer to Caravenna, den Hollander and Pétrélis (2012), pages 371–375.

The computation of the generating function G allows us to determine the exact value of β_c and to predict the behavior of the free energy close to criticality. However, the analytic expression of G is very complicated and only gives an indirect access to the free energy. Furthermore, this combinatorial method does not allow to study an observable which does not grow like L, for instance, inside the collapsed phase, the horizontal extension is of order $\sqrt{L}$ and this cannot be proven by such method.

A new approach has been developed in Nguyen and Pétrélis (2013) to work with the partition function directly. With the help of an algebraic manipulation of the Hamiltonian that will be described in Section 2.1, it is indeed possible to rewrite the partition function in (1.7) under the form

$$Z_{L,\beta} = c_\beta e^{\beta L} \sum_{N=1}^{L} (\Gamma(\beta))^N \mathbf{P}_\beta(\mathcal{V}_{N+1,L-N}), \tag{1.36}$$

where we recall (1.13) and (1.15) and where $\mathcal{V}_{n,k}$ is the set of those n-step trajectories of the random walk V whose geometric area $A_n = \sum_{i=1}^{n} |V_i|$ equals k, that is,

$$\mathcal{V}_{n,k} := \{(V_i)_{i=0}^{n} : A_n = k, V_n = 0\}. \tag{1.37}$$

Thus, the excess free energy $\widetilde{f}(\beta)$ is the exponential growth rate of the summation in (1.36). In this new expression of the partition function, the term indexed by

$N \in \{1, \dots, L\}$ in the summation corresponds to the contribution to the partition function of those trajectories $l \in \mathcal{L}_{N,L}$ (making N horizontal steps).

This new approach was used in Nguyen and Pétrélis (2013), Theorem 1.2, to derive a variational expression of the excess free energy, which allowed us to prove that the collapsed transition is second order with critical exponent $3/2$.

THEOREM 1.7 [Nguyen and Pétrélis (2013), Theorem 1.4]. *The phase transition is of order* $3/2$. *That is*, *there exist two constants* $c_1, c_2 > 0$ *such that for* ε *small enough*

$$c_1 \varepsilon^{3/2} \leq \widetilde{f}(\beta_c - \varepsilon) \leq c_2 \varepsilon^{3/2}. \tag{1.38}$$

With the present paper, we take the analysis of the phase transition two steps further (see Theorem B). In the first step, we establish the precise asymptotic: $\widetilde{f}(\beta_c - \varepsilon) \sim \gamma \varepsilon^{3/2}$ as $\varepsilon \searrow 0$ with γ an explicit constant. In the second step, we give an expression of γ in terms of the free energy of an auxiliary continuous model, that is, $\mathsf{F}_c = \lim_{T \to \infty} \frac{1}{T} \log \mathbf{E}[\exp(-\int_0^T |B(t)|\,dt)]$. Moreover, the Laplace transform of $\int_0^T |B(t)|\,dt$ was computed in Kac (1946) and this allows us to express F_c with the smallest zero (in modulus) of the derivative of the Airy function.

The question of the geometric conformation adopted by the polymer inside the collapsed phase has been raised and discussed by physicists in several papers, as for instance Brak et al. (1993). It was believed that the monomers arrange themselves in a succession of long vertical stretches of opposite directions that constitute large beads. In this paper, we prove with Theorem C, that the polymer makes *only one macroscopic bead* and that the number of monomers (located at the beginning and at the end of the polymer) which do not belong to this bead grows at most like $(\log L)^4$. We also make rigorous the conjecture concerning the horizontal extension of the polymer, since we identify the limit in probability of $N_L/\sqrt{L}$, which turns out to be the constant extracted from an optimization procedure. We also establish the convergence of properly rescaled lower and upper envelopes to a deterministic Wulff shape. In particular, the typical vertical displacement of the middle point, the $L/2$th monomer in a chain of length L, is of order $\sqrt{L}$.

There are numerical evidences that the vertical displacement of the endpoint grows as $L^{1/4}$ [see Brak et al. (1993), Table II, page 2394]. This turns out to be a consequence of the typical behavior of the fluctuations of the envelopes around the Wulff shape, and this is not the topic of the present paper.

Finally, let us stress the fact that the convergence, in the collapsed phase, to a deterministic Wulff shape (see Theorem E) comes from a fairly complex procedure that needs to establish three properties:

(i) The horizontal extension N_L is of order $\sqrt{L}$;
(ii) There is only one macroscopic bead;

(iii) When conditioned to be abnormally large, the geometric area of the associated V random walk ($\sum_i |V_i|$) is close to the modulus of its algebraic counterpart ($|\sum V_i|$).

There is no clear order in which to establish these properties and the proofs are intricate. For example, we need weak versions of (i) and (iii) to prove (ii) and then get a stronger version of (i).

2. Preparation: The main tools.

In this section, we introduce the three main tools that are used in this paper. In Section 2.1, we show how the partition function can be rewritten in terms of the random walk V of law $\mathbf{P}_\beta$ [recall (1.13)] and how studying this random walk under an appropriate conditioning can be used to derive some path properties under the polymer measure. In Section 2.2, we define the function $\delta \mapsto \mathfrak{h}_\beta(\delta)$ that appears in the expression of the excess free energy in Theorem A and we study its regularity. In Section 2.3, we consider the probability of some large deviations events under $\mathbf{P}_\beta$, and following Dobrushin and Hryniv (1996), we introduce an appropriate tilting under which the probability of such events decays only polynomially fast.

2.1. *Probabilistic representation of the partition function.*

In the first part of this section, we prove formula (1.36) and we show how the polymer measure can be expressed as the image measure by an appropriate transformation of the geometric random walk V introduced in (1.13). In the second part of the section, we focus on those trajectories that make only one bead and we show that, in terms of the auxiliary random walk V, these beads become excursions away from the origin.

2.1.1. *Auxiliary random walk.*

We display here the details of the proof of formula (1.36). Recall (1.4)–(1.7) and note that the $\tilde{\wedge}$ operator can be written as

$$x \,\tilde{\wedge}\, y = \big(|x| + |y| - |x+y|\big)/2 \qquad \forall x, y \in \mathbb{Z}. \tag{2.1}$$

Hence, for $\beta > 0$ and $L \in \mathbb{N}$, the partition function in (1.7) becomes

$$\begin{aligned} Z_{L,\beta} &= \sum_{N=1}^{L} \sum_{\substack{l \in \mathcal{L}_{N,L} \\ l_0 = l_{N+1} = 0}} \exp\left(\beta \sum_{n=1}^{N} |l_n| - \frac{\beta}{2} \sum_{n=0}^{N} |l_n + l_{n+1}|\right) \\ &= c_\beta e^{\beta L} \sum_{N=1}^{L} \left(\frac{c_\beta}{e^\beta}\right)^N \sum_{\substack{l \in \mathcal{L}_{N,L} \\ l_0 = l_{N+1} = 0}} \prod_{n=0}^{N} \frac{\exp(-(\beta/2)|l_n + l_{n+1}|)}{c_\beta}, \end{aligned} \tag{2.2}$$

where c_β was defined in (1.13). At this stage, we pick $N \in \{1, \dots, L\}$ and we introduce the one-to-one correspondence $T_N : \mathcal{V}_{N+1,L-N} \mapsto \mathcal{L}_{N,L}$ defined as $T_N(V)_i = (-1)^{i-1}V_i$ for all $i \in \{1,\dots N\}$. We pick $l \in \mathcal{L}_{N,L}$, we consider $V = (T_N)^{-1}(l)$ (see Figure 4) and we note that the increments $(U_i)_{i=1}^{N+1}$ of V necessarily satisfy $U_i := (-1)^{i-1}(l_{i-1}+l_i)$. Thus, the partition function in (2.2) becomes

$$Z_{L,\beta} = c_\beta e^{\beta L} \sum_{N=1}^{L} \left(\frac{c_\beta}{e^\beta}\right)^N \sum_{V \in \mathcal{V}_{N+1,L-N}} \mathbf{P}_\beta(V), \tag{2.3}$$

which immediately implies (1.36). A useful consequence of formula (2.3) is that, once conditioned on taking a given number of horizontal steps N, the polymer measure is exactly the image measure by the T_N-transformation of the geometric random walk V conditioned to return to the origin after $N+1$ steps and to make a geometric area $L-N$, that is,

$$P_{L,\beta}\big(l \in \cdot \,|\, N_L(l) = N\big) = \mathbf{P}_\beta\big(T_N(V) \in \cdot \,|\, V_{N+1} = 0, A_N = L-N\big). \tag{2.4}$$

2.1.2. *From beads to excursions*. We define Ω_L° as the subset of Ω_L containing those trajectories $l \in \Omega_L$ that have only one bead, that is, $n_L(l) = 1$. Thus, we can rewrite $\Omega_L^\circ := \bigcup_{N=1}^L \mathcal{L}_{N,L}^\circ$, where $\mathcal{L}_{N,L}^\circ$ is the subset of $\mathcal{L}_{N,L}$ defined as

$$\mathcal{L}_{N,L}^\circ = \big\{l \in \mathcal{L}_{N,L} : l_i \,\tilde{\curlywedge}\, l_{i+1} \neq 0 \ \forall i \in \{1,\dots,N-1\}\big\}, \tag{2.5}$$

and we denote by $Z_{L,\beta}^\circ$ the contribution to the partition function of those trajectories in Ω_L°, that is,

$$Z_{L,\beta}^\circ = \sum_{l \in \Omega_L^\circ} e^{\beta H_L(l)}. \tag{2.6}$$

We let also $\mathcal{V}_{n,k}^+$ be the subset containing those trajectories that return to the origin after n steps, satisfy $A_n = k$ and are strictly positive on $\{1,\dots,n\}$, that is,

$$\mathcal{V}_{n,k}^+ := \big\{V : V_n = 0, A_n = k, V_i > 0 \ \forall i \in \{1,\dots,n-1\}\big\}. \tag{2.7}$$

By mimicking (2.2) and by noticing that by the T_N-transformation, the subset $\mathcal{L}_{N,L}^\circ$ becomes $\mathcal{V}_{N+1,L-N}^+$ we obtain

$$Z_{L,\beta}^\circ = 2c_\beta e^{\beta L} \sum_{N=1}^{L} \big(\Gamma(\beta)\big)^N \mathbf{P}_\beta\big(\mathcal{V}_{N+1,L-N}^+\big). \tag{2.8}$$

2.2. *Construction and regularity of* $\mathfrak{h}_\beta$. We define the function $\mathfrak{h}_\beta$ in a slightly different way from (1.14), but we will see at the end of this section that the two definitions are equivalent. For $N \in \mathbb{N}$, $\delta \geq 0$, define

$$\mathfrak{h}_{N,\beta}(\delta) := \frac{1}{N} \log \mathbf{E}_\beta\big(e^{-\delta A_N} \mathbf{1}_{\{V_N=0\}}\big) \quad \text{and let} \quad \mathfrak{h}_\beta(\delta) = \lim_{N\to\infty} \mathfrak{h}_{N,\beta}(\delta). \tag{2.9}$$

LEMMA 2.1. (i) $\mathfrak{h}_\beta(\delta)$ *exists and is finite, nonpositive for all* $\beta > 0, \delta \geq 0$.
(ii) $\delta \mapsto \mathfrak{h}_\beta(\delta)$ *is continuous, convex and nonincreasing on* $[0, \infty)$.

PROOF. (i) For $N, M \in \mathbb{N}$, we restrict the partition of size $N + M$ to those trajectories that return to the origin at time N and use the Markov property to obtain

$$\mathbf{E}_\beta\big(e^{-\delta A_{N+M}}\mathbf{1}_{\{V_{N+M}=0\}}\big) \geq \mathbf{E}_\beta\big(e^{-\delta A_N}\mathbf{1}_{\{V_N=0\}}\big)\mathbf{E}_\beta\big(e^{-\delta A_M}\mathbf{1}_{\{V_M=0\}}\big). \tag{2.10}$$

Thus, $\{\log \mathbf{E}_\beta(e^{-\delta A_N}\mathbf{1}_{\{V_N=0\}})\}_{N\in\mathbb{N}}$ is a super-additive sequence that is bounded above by 0 and therefore the limit in (2.9) exists, is finite and satisfies

$$\mathfrak{h}_\beta(\delta) = \sup_{N\in\mathbb{N}} \frac{1}{N} \log \mathbf{E}_\beta\big(e^{-\delta A_N}\mathbf{1}_{\{V_N=0\}}\big) \leq 0. \tag{2.11}$$

(ii) The fact that $A_N \geq 0$ for all $N \in \mathbb{N}$ immediately entails that $\delta \mapsto \mathfrak{h}_\beta(\delta)$ is nonincreasing on $[0, \infty)$. By Hölder's inequality, the function $\delta \mapsto \mathfrak{h}_{N,\beta}(\delta)$ is convex for all $N \in \mathbb{N}$ and hence so is $\delta \mapsto \mathfrak{h}_\beta(\delta)$. Convexity and finiteness imply continuity on $(0, \infty)$. In order to prove the continuity at 0, we first note that $\lim_{\delta\to 0} \mathfrak{h}_\beta(\delta) = \sup_{\delta\geq 0} \mathfrak{h}_\beta(\delta)$. Then, with the help of formula (2.11) and via an exchange of suprema we obtain

$$\begin{aligned} \lim_{\delta\to 0} \mathfrak{h}_\beta(\delta) &= \sup_{\delta\geq 0} \sup_{N\in\mathbb{N}} \frac{1}{N} \log \mathbf{E}_\beta\big(e^{-\delta A_N}\mathbf{1}_{\{V_N=0\}}\big) \\ &= \sup_{N\in\mathbb{N}} \frac{1}{N} \log \mathbf{P}_\beta(V_N = 0) = 0. \end{aligned} \tag{2.12}$$

□

It remains to show that the two definitions of $\mathfrak{h}_\beta$ in (1.14) and (2.9) coincide. To that aim, it suffices to show that

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbf{E}_\beta\big(e^{-\delta A_N}\big) \leq \lim_{N\to\infty} \frac{1}{N} \log \mathbf{E}_\beta\big(e^{-\delta A_N}\mathbf{1}_{\{V_N=0\}}\big). \tag{2.13}$$

We set $\mathcal{I}_{N^2} := [-N^2, N^2] \cap \mathbb{Z}$ and we decompose $\mathbf{E}_\beta(e^{-\delta A_N})$ into the two partition functions $C_{N,\beta}$ and $B_{N,\beta}$ defined as

$$C_{N,\beta} = \mathbf{E}_\beta\big(e^{-\delta A_N}\mathbf{1}_{\{V_N\in\mathcal{I}_{N^2}\}}\big) \quad \text{and} \quad B_{N,\beta} = \mathbf{E}_\beta\big(e^{-\delta A_N}\mathbf{1}_{\{V_N\notin\mathcal{I}_{N^2}\}}\big). \tag{2.14}$$

Since $A_N \geq 0$ and since $\mathbf{E}_\beta[e^{\beta|U_1|/4}] < \infty$, Markov's inequality gives

$$B_{N,\beta} \leq \mathbf{E}_\beta[\mathbf{1}_{\{V_N\notin\mathcal{I}_{N^2}\}}] \leq \mathbf{P}_\beta\left(\sum_{i=1}^N |U_i| \geq N^2\right) \leq \frac{\mathbf{E}_\beta[e^{\beta|U_1|/4}]^N}{e^{(\beta/4)N^2}}, \tag{2.15}$$

which immediately implies that $\limsup_{N\to\infty} \frac{1}{N} \log B_{N,\beta} = -\infty$. Consequently,

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbf{E}_\beta\big(e^{-\delta A_N}\big) = \limsup_{N\to\infty} \frac{1}{N} \log C_{N,\beta}, \tag{2.16}$$

and since the cardinality of $\mathcal{I}_{N^2}$ grows polynomially, the proof of (2.13) will be complete once we show that

$$\limsup_{N\to\infty} \frac{1}{N} \log \sup_{x\in\mathcal{I}_{N^2}} \mathbf{E}_\beta\big(e^{-\delta A_N} \mathbf{1}_{\{V_N=x\}}\big) \le \lim_{N\to\infty} \frac{1}{N} \log \mathbf{E}_\beta\big(e^{-\delta A_N} \mathbf{1}_{\{V_N=0\}}\big). \tag{2.17}$$

For $x \in \mathbb{Z}$, we denote by $\mathbf{P}_{\beta,x}$ the law of $x + V$ where V is the random walk of law $\mathbf{P}_\beta$. We consider the partition function of size $2N$ and use Markov property at time N to obtain

$$\mathbf{E}_\beta\big(e^{-\delta A_{2N}} \mathbf{1}_{\{V_{2N}=0\}}\big) \ge \mathbf{E}_\beta\big(e^{-\delta A_N} \mathbf{1}_{\{V_N=x\}}\big) \mathbf{E}_{\beta,x}\big(e^{-\delta A_N} \mathbf{1}_{\{V_N=0\}}\big), \qquad x \in \mathbb{Z}. \tag{2.18}$$

By using the time reversal property of the random walk V, we can assert that $(V_N - V_{N-n}, 0 \le n \le N) \stackrel{d}{=} (V_n - V_0, 0 \le n \le N)$ and consequently, for all $x \in \mathbb{Z}$, it comes that

$$\begin{aligned} \mathbf{E}_{\beta,x}\big(e^{-\delta \sum_{n=1}^N |V_n|} \mathbf{1}_{\{V_N=0\}}\big) &= \mathbf{E}_\beta\big(e^{-\delta \sum_{n=1}^N |V_n+x|} \mathbf{1}_{\{V_N=-x\}}\big) \\ &= \mathbf{E}_\beta\big(e^{-\delta \sum_{n=1}^N |V_N - V_{N-n}+x|} \mathbf{1}_{\{V_N=-x\}}\big) \\ &= \mathbf{E}_\beta\big(e^{-\delta \sum_{n=1}^{N-1} |V_n|} \mathbf{1}_{\{V_N=-x\}}\big). \end{aligned} \tag{2.19}$$

Thanks to the symmetry of V and since $\sum_{n=1}^{N-1} |V_n| \le A_N$, the inequalities (2.18) and (2.19) allow us to write

$$\mathbf{E}_\beta\big(e^{-\delta A_{2N}} \mathbf{1}_{\{V_{2N}=0\}}\big) \ge \Big[\sup_{x\in\mathcal{I}_{N^2}} \mathbf{E}_\beta\big(e^{-\delta A_N} \mathbf{1}_{\{V_N=x\}}\big)\Big]^2. \tag{2.20}$$

It remains to apply $\frac{1}{2N}\log$ in both sides of (2.20) and to let $N \to \infty$ to obtain (2.17), which completes the proof.

2.3. *Large deviation estimates*. In this section, we introduce the techniques that will be required to estimate the probability of some large deviation events associated with trajectories making a large arithmetic area. Such estimates will be needed in Section 4 to approximate the probability that, under the polymer measure, the trajectories make only one bead.

Following Dobrushin and Hryniv [Dobrushin and Hryniv (1996)], for $n \in \mathbb{N}$, we define

$$Y_n := \frac{1}{n}(V_0 + V_1 + \cdots + V_{n-1}), \tag{2.21}$$

and for a given $q \in (0, \infty) \cap \frac{\mathbb{N}}{n}$, we focus on both probabilities $\mathbf{P}_\beta(Y_n = nq, V_n = 0)$ and $\mathbf{P}_\beta(Y_n = nq, V_n = 0, V_i > 0\ \forall i \in \{1, \dots, n-1\})$. Our aim is to identify the exponential rate at which such probabilities are decreasing and their asymptotic polynomial correction. To that aim, we will use an *exponential tilting* of the probability measure $\mathbf{P}_\beta$ (through the Cramér transform) in combination with a local limit theorem. Under the tilted probability measure, the event $\{Y_n = nq, V_n = 0\}$ is not of large deviation type anymore since its probability decays at polynomial speed instead of exponential speed, as will be seen in Section 6.

For the ease of notation, we set $\Lambda_n := (Y_n, V_n)$ and we denote its logarithmic moment generating function by $\mathfrak{L}_{\Lambda_n}(\mathbf{h})$ for $\mathbf{h} := (h_0, h_1) \in \mathbb{R}^2$, that is,

$$\mathfrak{L}_{\Lambda_n}(\mathbf{h}) := \log \mathbf{E}_\beta\left[e^{h_0 Y_n + h_1 V_n}\right] = \sum_{i=1}^{n} \mathfrak{L}\left(\left(1 - \frac{i}{n}\right)h_0 + h_1\right). \tag{2.22}$$

Clearly, $\mathfrak{L}_{\Lambda_n}(\mathbf{h})$ is finite for all $\mathbf{h} \in \mathcal{D}_n$ with

$$\mathcal{D}_n := \left\{(h_0, h_1) \in \mathbb{R}^2 : h_1 \in \left(-\frac{\beta}{2}, \frac{\beta}{2}\right), \left(1 - \frac{1}{n}\right)h_0 + h_1 \in \left(-\frac{\beta}{2}, \frac{\beta}{2}\right)\right\}. \tag{2.23}$$

With the help of (2.22) and for $\mathbf{h} = (h_0, h_1) \in \mathcal{D}_n$, we define the $\mathbf{h}$-tilted distribution by

$$\frac{\mathrm{d}\mathbf{P}_{n,\mathbf{h}}}{\mathrm{d}\mathbf{P}_\beta}(V) = e^{h_0 Y_n + h_1 V_n - \mathfrak{L}_{\Lambda_n}(H)}. \tag{2.24}$$

For a given $n \in \mathbb{N}$ and $q \in \frac{\mathbb{N}}{n}$, the exponential tilt is given by $\mathbf{h}_n^q := (h_{n,0}^q, h_{n,1}^q)$ which, by Lemma 5.4 in Section 5.1, is the unique solution of

$$\mathbf{E}_{n,\mathbf{h}}\left(\frac{\Lambda_n}{n}\right) = \nabla\left[\frac{1}{n}\mathfrak{L}_{\Lambda_n}\right](\mathbf{h}) = (q, 0) \tag{2.25}$$

and, therefore, we have the equality

$$\mathbf{P}_\beta\big(\Lambda_n = (nq, 0)\big) = \mathbf{P}_{n,\mathbf{h}_n^q}\big(\Lambda_n = (nq, 0)\big)e^{n(-h_{n,0}^q q + (1/n)\mathfrak{L}_{\Lambda_n}(\mathbf{h}_n^q))}. \tag{2.26}$$

From (2.26), it is easy to deduce that the exponential decay rate of $\mathbf{P}_\beta(\Lambda_n = (nq, 0))$ is given by the quantity $-h_{n,0}^q q + \frac{1}{n}\mathfrak{L}_{\Lambda_n}(\mathbf{h}_n^q)$ and that the polynomial correction is associated with $\mathbf{P}_{n,\mathbf{h}_n^q}(\Lambda_n = (nq, 0))$. To be more specific, we first state a proposition which gives a local central limit theorem for the tilted law $\mathbf{P}_{n,\mathbf{h}_n^q}$.

PROPOSITION 2.2. *For $[q_1, q_2] \subset (0, \infty)$, there exist $C > 0, n_0 > 0$ such that for all*[2] *$q \in [q_1, q_2]$ and $n \geq n_0$ we have*

$$\frac{1}{Cn^2} \leq \mathbf{P}_{n,\mathbf{h}_n^q}(Y_n = nq, V_n = 0) \leq \frac{C}{n^2}. \tag{2.27}$$

[2]To be thorough, we should restrict ourselves to q such that $n^2 q \in \mathbb{N}$. To ease notation, we shall omit this restriction in the sequel.

The following proposition shows that the exponential decay rate induced by the change of probability in (2.24) can be controlled uniformly in n.

PROPOSITION 2.3 (Decay rate of large area probability). *For* $[q_1, q_2] \subset (0, +\infty)$, *there exist* $c_1, c_2 > 0$ *and* $n_0 \in \mathbb{N}$ *such that*

$$\left\| \left[\frac{1}{n} \mathfrak{L}_{\Lambda_n}(\mathbf{h}_n^q) - h_{n,0}^q q \right] - \left[\mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q,0)) - \widetilde{h}_0(q,0) q \right] \right\| \le \frac{c_1}{n} \quad \text{for } n \ge n_0, q \in [q_1, q_2] \tag{2.28}$$

and

$$\|\mathbf{h}_n^q - \widetilde{\mathbf{h}}(q,0)\| \le \frac{c_2}{n} \qquad \text{for } n \ge n_0, q \in [q_1, q_2]. \tag{2.29}$$

Propositions 2.2 and 2.3 will be proven in Sections 6 and 5.1, respectively. With the help of (2.26) and by applying Propositions 2.2 and 2.3 we can finally give some sharp upper and lower bounds of $\mathbf{P}_\beta(Y_n = nq, V_n = 0)$.

PROPOSITION 2.4. *For* $[q_1, q_2] \subset (0, \infty)$, *there exist* $C_1 > C_2 > 0$ *and* $n_0 \in \mathbb{N}$ *such that for all* $q \in [q_1, q_2]$ *and* $n \ge n_0$ *we have*

$$\begin{aligned} \frac{C_2}{n^2} e^{n[-\widetilde{h}_0(q,0)q + \mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q,0))]} &\le \mathbf{P}_\beta(Y_n = nq, V_n = 0) \\ &\le \frac{C_1}{n^2} e^{n[-\widetilde{h}_0(q,0)q + \mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q,0))]}. \end{aligned} \tag{2.30}$$

In addition, we shall need in this paper a precise lower bound on the probability that, under $\mathbf{P}_\beta$, the random walk V makes only one excursion away from the origin, conditionally on having a large prescribed area. To our knowledge, such an estimate is not available in the existing literature. Recall the definition of Y_n in (2.21).

PROPOSITION 2.5 (Unique excursion for large area). *For* $[q_1, q_2] \subset (0, \infty)$, *there exist* $C > 0$, $\mu > 0$ *and* $n_0 \in \mathbb{N}$ *such that for all* $q \in [q_1, q_2]$ *and every* $n \ge n_0$

$$\mathbf{P}_\beta(V_i > 0, 0 < i < n \mid Y_n = nq, V_n = 0) \ge \frac{C}{n^\mu}. \tag{2.31}$$

Although we can show that for the tilted law $\mathbf{P}_{n,\mathbf{h}_n^q}$ (thanks to the positive, resp., negative drifts of the increments close to 0, resp., close to n) there exists a $C(q_1, q_2) > 0$ so that for $q \in [q_1, q_2]$ and n large enough

$$\mathbf{P}_{n,\mathbf{h}_n^q}(V_i > 0, 0 < i < n \mid V_n = 0) > C(q_1, q_2),$$

and although we think that a similar result holds true for the LHS in (2.31), we are unable to handle the conditioning by $Y_n = nq$ satisfactorily.

3. The order of the phase transition. In Section 3.1 below, we prove Theorem A that expresses the excess free energy as the solution of an equation involving the function $\mathfrak{h}_\beta$ introduced in Section 2.2. In Section 3.2, we first state Lemma 3.1 which provides the behavior of $\mathfrak{h}_\beta(\widetilde{f}(\beta))$ close to β_c and then we combine this lemma with Theorem A to complete the proof of Theorem B. Finally, in Section 3.3 we give a proof of Lemma 3.1.

3.1. *Proof of Theorem* A (*Free energy equation*). By the representation formula (1.36) and the definition of $\widetilde{f}$, we have $\widetilde{f}(\beta) = \lim_{L\to\infty} \frac{1}{L} \log \widetilde{Z}_{L,\beta}$, where

$$\widetilde{Z}_{L,\beta} := \sum_{N=1}^{L} \bigl(\Gamma(\beta)\bigr)^N \mathbf{P}_\beta(\mathcal{V}_{N+1,L-N}). \tag{3.1}$$

As a consequence, the excess free energy satisfies $\widetilde{f}(\beta) = -\log R$ where R is the radius of convergence of the generating function $G(z) = \sum_{L=1}^{\infty} \widetilde{Z}_{L,\beta} z^L$, that is,

$$\widetilde{f}(\beta) = \sup\left\{\delta \ge 0 \colon \sum_{L=1}^{\infty} \widetilde{Z}_{L,\beta} e^{-\delta L} = +\infty\right\}, \tag{3.2}$$

if the set is nonempty and $\widetilde{f}(\beta) = 0$ otherwise. We recall (1.37) and we use (3.1) to rewrite the sum in (3.2) as

$$\begin{aligned}
\sum_{L=1}^{\infty} \widetilde{Z}_{L,\beta} e^{-\delta L} &= \sum_{L=1}^{\infty} \sum_{N=1}^{L} \bigl(\Gamma(\beta) e^{-\delta}\bigr)^N \sum_{\substack{V_0 = V_{N+1} = 0 \\ A_N = L-N}} \mathbf{P}_\beta(V) e^{-\delta(L-N)} \\
&= \sum_{L=1}^{\infty} \sum_{N=1}^{L} \bigl(\Gamma(\beta) e^{-\delta}\bigr)^N \mathbf{E}_\beta\bigl(e^{-\delta A_N} \mathbf{1}_{\{A_N = L-N, V_{N+1}=0\}}\bigr) \\
&= \sum_{N=1}^{\infty} \bigl(\Gamma(\beta) e^{-\delta}\bigr)^N \mathbf{E}_\beta\bigl(e^{-\delta A_N} \mathbf{1}_{\{V_{N+1}=0\}}\bigr).
\end{aligned} \tag{3.3}$$

Since $A_N = A_{N+1}$ on the set $\{V_{N+1} = 0\}$ and by using the definition of $\mathfrak{h}_{N,\beta}(\delta)$ in (2.9), the equality (3.3) becomes

$$\sum_{L=1}^{\infty} \widetilde{Z}_{L,\beta} e^{-\delta L} = \sum_{N=1}^{\infty} \exp\left(N\left[\log \Gamma(\beta) - \delta + \frac{N+1}{N} \mathfrak{h}_{N+1,\beta}(\delta)\right]\right), \tag{3.4}$$

which together with (3.2) gives $\widetilde{f}(\beta) = \sup\{\delta \ge 0 \colon \log \Gamma(\beta) - \delta + \mathfrak{h}_\beta(\delta) > 0\}$. Since $\mathfrak{h}_\beta(\delta) \le 0$, it follows that $\widetilde{f}(\beta) = 0$ if $\Gamma(\beta) \le 1$. When $\Gamma(\beta) > 1$, Lemma 2.1 gives that $\delta \mapsto -\delta + \mathfrak{h}_\beta(\delta)$ is continuous, decreasing, nonpositive on $[0, \infty)$, equals 0 at $\delta = 0$ and tends to $-\infty$ when $\delta \to \infty$. Therefore, $\widetilde{f}(\beta) > 0$ and is the unique solution of the equation $\log\Gamma(\beta)-\delta+\mathfrak{h}_\beta(\delta)=0$. In addition, by recalling the definition of the collapsed phase (1.11) and the extended phase (1.12), we can observe that

$$\mathcal{C}=\{\beta:\Gamma(\beta)\le 1\}\quad\text{and}\quad\mathcal{E}=\{\beta:\Gamma(\beta)>1\}. \tag{3.5}$$

We note that $\beta\mapsto\Gamma(\beta)$ is decreasing on $[0,\infty)$ [recall (1.13) and (1.15)] and therefore, the collapse transition occurs at β_c, the unique positive solution of the equation $\Gamma(\beta)=1$.

3.2. *Proof of Theorem* B (*Phase transition asymptotics*). We display here the proof of Theorem B subject to Lemma 3.1 below, that will be proven in Section 3.3 afterward.

LEMMA 3.1.

$$\lim_{\beta\to\beta_c}\frac{\mathfrak{h}_\beta(\widetilde{f}(\beta))}{\widetilde{f}(\beta)^{2/3}}=-\varsigma_2, \tag{3.6}$$

where we recall that ς_2 *was defined in* (1.18).

Our aim is to study the asymptotic behavior of the equation in Theorem A near the critical point. We recall (1.15) and we perform a first-order Taylor expansion of $\Gamma(\beta)$ near β_c which gives $\log\Gamma(\beta_c-\varepsilon)=\varsigma_1\varepsilon(1+o(1))$ as $\varepsilon\searrow 0$. Next, we consider the function $\mathfrak{h}_\beta$ near β_c and it follows from Lemma 3.1 that when $\varepsilon\searrow 0$

$$h_{\beta_c-\varepsilon}(\widetilde{f}(\beta_c-\varepsilon))=-\varsigma_2\widetilde{f}(\beta_c-\varepsilon)^{2/3}(1+o(1)). \tag{3.7}$$

Therefore, by plugging (3.7) and the expansion of $\log\Gamma(\beta_c-\varepsilon)$ in the equation in Theorem A that is verified by the excess free energy, we obtain that

$$\varsigma_1\varepsilon(1+o(1))-\widetilde{f}(\beta_c-\varepsilon)-\varsigma_2\widetilde{f}(\beta_c-\varepsilon)^{2/3}(1+o(1))=0, \tag{3.8}$$

which allows to conclude that

$$\widetilde{f}(\beta_c-\varepsilon)\sim\left(\frac{\varsigma_1}{\varsigma_2}\right)^{3/2}\varepsilon^{3/2}\qquad\text{as }\varepsilon\searrow 0, \tag{3.9}$$

and the proof is complete.

3.3. *Asymptotics of* $\mathfrak{h}_\beta$.

3.3.1. *Heuristics*. Let us give the heuristic explanation of why $\mathfrak{h}_\beta(\delta)\sim -c\delta^{2/3}$ for some constant $c>0$. The main idea is to decompose the trajectory of the random walk V into independent blocks of length $T\delta^{-2/3}$ for $T\in\mathbb{N}$ and δ small enough: we have approximately $N/(T\delta^{-2/3})$ such blocks. Hence, as $\delta\searrow 0$, we can estimate

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbf{E}_\beta(e^{-\delta A_N})\sim\lim_{T\to\infty}\frac{\delta^{2/3}}{T}\log\mathbf{E}_\beta(e^{-\delta A_{T\delta^{-2/3}}}). \tag{3.10}$$

It is well known that for such random walks (assume that $\mathbf{E}_\beta(U_1^2)=1$) [see Durrett (2010), page 405]

$$k^{-3/2}\sum_{i=1}^{Tk}|V_i| \to \mathcal{L}\int_0^T |B(t)|\,dt \qquad \text{as } k\to\infty, \tag{3.11}$$

where B is a standard Brownian motion. Now, let $k=\delta^{-2/3}$ and since $|e^{-\delta A_{T\delta^{-2/3}}}|\le 1$, we conclude that

$$\mathbf{E}_\beta\big(e^{-\delta A_{T\delta^{-2/3}}}\big) \to \mathbf{E}\big(e^{-\int_0^T |B(t)|\,dt}\big) \qquad \text{as } \delta\to 0. \tag{3.12}$$

This convergence and (3.10) would immediately imply $\mathfrak{h}_\beta(\delta)\sim -c\delta^{2/3}$ where c can be estimated via the distribution of the *Brownian area*, that is,

$$c = -\lim_{T\to\infty}\frac{1}{T}\log \mathbf{E}\big(e^{-\int_0^T |B(t)|\,dt}\big) > 0. \tag{3.13}$$

3.3.2. *Proof of Lemma* 3.1.

3.3.2.1. *Upper bound*. Pick $T\in\mathbb{N}$, $\delta>0$ such that $\delta^{-2/3}\in\mathbb{N}$ and let $\Delta := \delta^{-2/3}$. We take N that satisfies $N/(T\Delta)\in\mathbb{N}$ and partition $\{1,\dots,N\}$ into $k=N/(T\Delta)$ intervals of length $T\Delta$. By the Markov property of V, we decompose $\mathbf{E}_\beta(e^{-\delta A_N})$ with respect to the position occupied by the random walk V at times $T\Delta, 2T\Delta,\dots,(k-1)T\Delta$,

$$\begin{aligned}\mathbf{E}_\beta\big(e^{-\delta A_N}\big) &= \sum_{\substack{x_0=0,x_i\in\mathbb{Z}\\ i=1,\dots,k}}\prod_{i=0}^{k-1}\mathbf{E}_{\beta,x_i}\big(e^{-\delta A_{T\Delta}}\mathbf{1}_{\{V_{T\Delta}=x_{i+1}\}}\big)\\ &\le \Big[\sup_{x\in\mathbb{Z}}\mathbf{E}_{\beta,x}\big(e^{-\delta A_{T\Delta}}\big)\Big]^k.\end{aligned} \tag{3.14}$$

With the help of Lemma 3.2 below, we can replace the supremum in the RHS of (3.14) by the term indexed by $x=0$ only. The proof of Lemma 3.2 is postponed to Appendix B.

LEMMA 3.2. *For all $\delta>0, n\in\mathbb{N}$ and $x,x'\in\mathbb{Z}$ such that $|x'|\ge|x|$, the following inequality holds true*

$$\mathbf{E}_{\beta,x'}\big(e^{-\delta A_n}\big)\le \mathbf{E}_{\beta,x}\big(e^{-\delta A_n}\big). \tag{3.15}$$

Therefore, (3.14) becomes

$$\mathbf{E}_\beta\big(e^{-\delta A_N}\big)\le \big[\mathbf{E}_\beta\big(e^{-\delta A_{T\Delta}}\big)\big]^{N/(T\Delta)}. \tag{3.16}$$

Recall that $\Delta:=\delta^{-2/3}$, apply $\frac{1}{N}\log$ to both sides of (3.16) and let $N\to\infty$ to obtain, for $\beta>0$ and $\delta>0$, that

$$\frac{\mathfrak{h}_\beta(\delta)}{\delta^{2/3}}\le \frac{1}{T}\log \mathbf{E}_\beta\big(e^{-\delta A_{T\Delta}}\big). \tag{3.17}$$

In what follows, we need a uniform version (in β) of the convergence of $\mathbf{E}_\beta(e^{-\delta A_{T\Delta}})$ toward $\mathbf{E}(e^{-\int_0^T |B(t)|\,dt})$ as $\delta \to 0$. For this reason, we introduce the strong approximation theorem [Sakhanenko (1980)] to approximate the partial sums of independent random variables U in the RHS in (3.17) by independent normal random variables.

THEOREM 3.3 [Shao (1995), Theorem B]. *Denote by σ_β^2 the variance of the random variable U_1 under $\mathbf{P}_\beta$. We can redefine $\{U_i, i \geq 1\}$ (denoted by U^β) on a richer probability space together with a sequence of independent standard normal random variables $\{X_i, i \geq 1\}$ such that for every $p > 2$, $x > 0$,*

$$\mathbf{P}\left(\max_{i \leq n}\left|\sum_{j=1}^{i} U_j^\beta - \sigma_\beta \sum_{j=1}^{i} X_j\right| \geq x\right) \leq (Ap)^p x^{-p} \sum_{i=1}^{n} \mathbf{E}|U_i^\beta|^p, \tag{3.18}$$

where A is an absolute positive constant.

We let also, for $n \in \mathbb{N}$, $Y_n = \sum_{i=1}^n X_i$, $A_n(Y) = \sum_{i=1}^n |Y_i|$ and redefine $V_n^\beta = \sum_{i=1}^n U_i^\beta$, $A_n(V^\beta) = \sum_{i=1}^n |V_i^\beta|$. We pick $T > 0$, $p > 2$, $\vartheta > 0$ and K a compact subset of $(0, \infty)$. We use Theorem 3.3 and the fact that [recall (1.13)] $\mathbf{E}[|U_1^\beta|^p]$ is bounded from above uniformly in $\beta \in K$, to assert that there exists a constant $c_{p,K} > 0$ such that for all $\Delta > 0$ and $\beta \in K$

$$\mathbf{P}\Big(\max_{i \leq T\Delta}|V_i^\beta - \sigma_\beta Y_i| \geq \Delta^\vartheta\Big) \leq c_{p,K} T \Delta^{1-\vartheta p}. \tag{3.19}$$

Note that on the event $\{\max_{i \leq T\Delta}|V_i^\beta - \sigma_\beta Y_i| < \Delta^\vartheta\}$, we obviously have $|A_{T\Delta}(V^\beta) - \sigma_\beta A_{T\Delta}(Y)| \leq T\Delta^{\vartheta+1}$. Therefore, since $x \mapsto \exp(-x)$ is 1-Lipschitz on $[0, \infty)$ and since $\Delta = \delta^{-2/3}$, we can write that for $\beta \in K$ and $\delta > 0$

$$\begin{aligned} &\big|\mathbf{E}\big(e^{-\delta A_{T\Delta}(V^\beta)} - e^{-\delta\sigma_\beta A_{T\Delta}(Y)}\big)\big| \\ &\quad \leq \mathbf{P}\Big(\max_{i \leq T\Delta}|V_i^\beta - \sigma_\beta Y_i| \geq \Delta^\vartheta\Big) + \delta T \Delta^{\vartheta+1} \\ &\quad \leq c_{p,K} T \delta^{(2/3)(\vartheta p - 1)} + T \delta^{(1/3)(1-2\vartheta)}. \end{aligned} \tag{3.20}$$

We chose $p = 3$ and $\vartheta \in (1/3, 1/2)$ and plug it in the RHS of (3.17) to obtain that for $\beta \in K$ and $\delta > 0$,

$$\frac{\mathfrak{h}_\beta(\delta)}{\delta^{2/3}} \leq \frac{1}{T}\log\big[\mathbf{E}\big(e^{-\delta\sigma_\beta A_{T\Delta}(Y)}\big) + c_{3,K} T \delta^{2(3\vartheta-1)/3} + T\delta^{(1-2\vartheta)/3}\big]. \tag{3.21}$$

LEMMA 3.4. *Let K be a compact subset of $(0, +\infty)$. For $T > 0$ and $\varepsilon > 0$ there exists a $\delta_0 > 0$ such that for $\delta \leq \delta_0$ (with $\Delta = \delta^{-2/3}$),*

$$\sup_{\beta \in K}\big|\mathbf{E}\big(e^{-\delta\sigma_\beta A_{T\Delta}(Y)}\big) - \mathbf{E}\big(e^{-\sigma_\beta \int_0^T |B(t)|\,dt}\big)\big| < \varepsilon, \tag{3.22}$$

where B is a standard Brownian motion.

PROOF. We can consider $\{B(t), t \geq 0\}$ and $\{y_i, i \geq 1\}$ on the same probability space by letting $y_i = B(i) - B(i-1)$, and thus $Y_i := y_1 + \cdots + y_i = B(i)$ for $i \in \mathbb{N}$. We recall that $A_{T\Delta}(Y) = \sum_{i=1}^{T\Delta} |B(i)|$ and therefore, by Brownian scaling we note that

$$\Delta^{-3/2} A_{T\Delta}(Y) \stackrel{d}{=} \Delta^{-1} \sum_{i=1}^{T\Delta} |B(i/\Delta)|.$$

Consequently, by recalling that $\delta = \Delta^{-3/2}$ we can replace $\mathbf{E}(e^{-\delta\sigma_\beta A_{T\Delta}(Y)})$ in the LHS of (3.22) by $\mathbf{E}(e^{-\sigma_\beta \Delta^{-1} \sum_{i=1}^{T\Delta} |B(i/\Delta)|})$. Since the exponential function is 1-Lipschitz on $(-\infty, 0]$, we have

$$\begin{aligned}\sup_{\beta \in K} &\big|\mathbf{E}\big(e^{-\sigma_\beta \Delta^{-1} \sum_{i=1}^{T\Delta} |B(i/\Delta)|}\big) - \mathbf{E}\big(e^{-\sigma_\beta \int_0^T |B(t)|\,dt}\big)\big| \\ &\leq \max_{\beta \in K}\{\sigma_\beta\} \mathbf{E}\left[\left|\Delta^{-1} \sum_{i=1}^{T\Delta} |B(i/\Delta)| - \int_0^T |B(t)|\,dt\right|\right].\end{aligned}$$

Since $\max_{\beta \in K}\{\sigma_\beta\} < \infty$, since by Riemann sum approximation we know that

$$\Delta^{-1} \sum_{i=1}^{T\Delta} |B(i/\Delta)| \xrightarrow[\Delta \to \infty]{\text{a.s.}} \int_0^T |B(t)|\,dt. \tag{3.23}$$

It is easy to see that

$$\sup_{\Delta > 0} \mathbf{E}\left(\Delta^{-1} \sum_{i=1}^{T\Delta} |B(i/\Delta)|^2\right) < \infty,$$

and this implies uniform integrability which, combined with the almost sure convergence implies the convergence in L^1

$$\lim_{\Delta \to \infty} \mathbf{E}\left[\left|\Delta^{-1} \sum_{i=1}^{T\Delta} |B(i/\Delta)| - \int_0^T |B(t)|\,dt\right|\right] = 0. \tag{3.24}$$

This completes the proof. □

We resume the proof of the upper bound. Since $\vartheta \in (1/3, 1/2)$, the RHS of (3.20) vanishes as $\delta \to 0$ uniformly in $\beta \in K$. Thus, we can replace δ by $\widetilde{f}(\beta_c)$ in (3.21) and use Lemma 3.4 and the fact that $\lim_{\varepsilon \to 0^+} \widetilde{f}(\beta_c - \varepsilon) = 0$ to conclude that, for all $T > 0$,

$$\limsup_{\varepsilon \to 0^+} \frac{\mathfrak{h}_\beta(\widetilde{f}(\beta_c - \varepsilon))}{\widetilde{f}(\beta_c - \varepsilon)^{2/3}} \leq \frac{1}{T} \log \mathbf{E}\big(e^{-\sigma_{\beta_c} \int_0^T |B(t)|\,dt}\big). \tag{3.25}$$

It remains to let T tend to infinity and to recall (1.18) to obtain

$$\limsup_{\varepsilon \to 0^+} \frac{\mathfrak{h}_\beta(\widetilde{f}(\beta_c - \varepsilon))}{\widetilde{f}(\beta_c - \varepsilon)^{2/3}} \leq -\varsigma_2. \tag{3.26}$$

3.3.2.2. *Lower bound.* Recall that $T \in \mathbb{N}$, $\delta > 0$ and $\Delta = \delta^{-2/3} \in \mathbb{N}$. We also take $N \in \mathbb{N}$ such that $N/(T\Delta) \in \mathbb{N}$. Pick $\eta > 0$ and use the decomposition in (3.14) to obtain

$$\mathbf{E}_\beta(e^{-\delta A_N}) \geq \sum_{\substack{x_0=0, x_i \in [-\eta\sqrt{\Delta}, \eta\sqrt{\Delta}] \\ i=1,\dots,k}} \prod_{i=0}^{k-1} \mathbf{E}_{\beta,x_i}(e^{-\delta A_{T\Delta}} \mathbf{1}_{\{V_{T\Delta}=x_{i+1}\}}) \tag{3.27}$$

$$\geq \Big[\inf_{x\in[-\eta\sqrt{\Delta},\eta\sqrt{\Delta}]} \mathbf{E}_{\beta,x}(e^{-\delta A_{T\Delta}} \mathbf{1}_{\{V_{T\Delta}\in[-\eta\sqrt{\Delta},\eta\sqrt{\Delta}]\}})\Big]^{N/(T\Delta)}. \tag{3.28}$$

For any integer $x \in [-\eta\sqrt{\Delta}, \eta\sqrt{\Delta}]$, we consider the two sets of paths

$$\Pi_1^x = \{(V_i)_{i=0}^{T\Delta} : V_0 = x,\ V_{T\Delta} \in [-\eta\sqrt{\Delta}, \eta\sqrt{\Delta}]\} \tag{3.29}$$

and

$$\Pi_2 = \{(V_i)_{i=0}^{T\Delta} : V_0 = 0,\ V_{T\Delta} \in [-\eta\sqrt{\Delta}, 0]\}. \tag{3.30}$$

Clearly, if $V = (V_i)_{i=0}^{T\Delta} \in \Pi_2$, then the trajectory $V + x$ starts at $x \in [0, \eta\sqrt{\Delta}]$ and is an element of Π_1^x. Similarly, for $x \in [-\eta\sqrt{\Delta}, 0]$, $\Pi_2' + x \subseteq \Pi_1^x$ where

$$\Pi_2' = \{(V_i)_{i=0}^{T\Delta} : V_0 = 0,\ V_{T\Delta} \in [0, \eta\sqrt{\Delta}]\}. \tag{3.31}$$

Since $\mathbf{P}_\beta(V \in \Pi_2) = \mathbf{P}_\beta(V \in \Pi_2')$, we conclude that

$$\mathbf{P}_{\beta,x}(V \in \Pi_1^x) \geq \mathbf{P}_\beta(V \in \Pi_2') \qquad \text{for all } x \in [-\eta\sqrt{\Delta}, \eta\sqrt{\Delta}]. \tag{3.32}$$

Moreover, for any $V^\star \in \Pi_1^x$,

$$\delta \sum_{i=1}^{T\Delta} |V_i^\star| = \delta \sum_{i=1}^{T\Delta} |x + V_i| \leq \delta \sum_{i=1}^{T\Delta} |V_i| + \delta T\Delta|x| \leq \delta \sum_{i=1}^{T\Delta} |V_i| + \eta T, \tag{3.33}$$

where the trajectory V satisfies $V_0 = 0$. Combining (3.32) and (3.33), we then have, for $x \in [-\eta\sqrt{\Delta}, \eta\sqrt{\Delta}]$,

$$\mathbf{E}_{\beta,x}(e^{-\delta A_{T\Delta}} \mathbf{1}_{\{V_{T\Delta}\in[-\eta\sqrt{\Delta},\eta\sqrt{\Delta}]\}}) \geq e^{-\eta T} \mathbf{E}_\beta(e^{-\delta A_{T\Delta}} \mathbf{1}_{\{V_{T\Delta}\in[0,\eta\sqrt{\Delta}]\}}). \tag{3.34}$$

By plugging the lower bound above into (3.27) and by using the symmetry of V we immediately get

$$\mathbf{E}_\beta(e^{-\delta A_N}) \geq [e^{-\eta T} \mathbf{E}_\beta(e^{-\delta A_{T\Delta}} \mathbf{1}_{\{V_{T\Delta}\in[0,\eta\sqrt{\Delta}]\}})]^{N/T\Delta}, \tag{3.35}$$

which, by applying $\frac{1}{N}\log$ to both sides in (3.35) and by letting $N \to \infty$, gives, for all $\beta > 0$,

$$\frac{\mathfrak{h}_\beta(\delta)}{\delta^{2/3}} \geq \frac{1}{T} \log \mathbf{E}_\beta(e^{-\delta A_{T\Delta}} \mathbf{1}_{\{V_{T\Delta}\in[0,\eta\sqrt{\Delta}]\}}) - \eta, \qquad \delta, \eta > 0. \tag{3.36}$$

At this stage, we proceed as in the upper bound [from (3.17)] to obtain, for all $T \in \mathbb{N}, \eta > 0$,

$$\liminf_{\beta \to \beta_c} \frac{\mathfrak{h}_\beta(\widetilde{f}(\beta))}{\widetilde{f}(\beta)^{2/3}} \geq \frac{1}{T} \log \mathbf{E}\big(e^{-\sigma_{\beta_c} \int_0^T |B(t)|\,dt} \mathbf{1}_{\{B(T) \in [0,\eta]\}}\big) - \eta. \tag{3.37}$$

It remains to show that for all $\eta > 0$ we have

$$\begin{aligned} &\lim_{T \to \infty} \frac{1}{T} \log \mathbf{E}\big(e^{-\sigma_{\beta_c} \int_0^T |B(t)|\,dt} \mathbf{1}_{\{B(T) \in [0,\eta]\}}\big) \\ &\quad = \lim_{T \to \infty} \frac{1}{T} \log \mathbf{E}\big(e^{-\sigma_{\beta_c} \int_0^T |B(t)|\,dt}\big), \end{aligned} \tag{3.38}$$

but the latter convergence can be obtained by adapting the proof of (2.13) to the continuous setting and for conciseness we will not give the details of the proof here. Then, by recalling (1.18), we achieve the bound

$$\liminf_{\beta \to \beta_c} \frac{\mathfrak{h}_\beta(\widetilde{f}(\beta))}{\widetilde{f}(\beta)^{2/3}} \geq -\varsigma_2 - \eta, \tag{3.39}$$

for all $\eta > 0$. It remains to let $\eta \to 0$ to complete the proof.

4. Geometry of the collapsed phase. In Section 4.1 below, a proof of Theorem C is displayed subject to Lemma 4.1, which ensures that the horizontal extension of the polymer inside the collapsed phase is of order $\sqrt{L}$, and to Proposition 4.2, which provides a sharp estimate of the partition function restricted to those trajectories making only one bead. Proposition 4.2 is proven in Section 4.2 subject to Lemma 4.4, which is the counterpart of Lemma 4.1 for the one bead trajectory and to Proposition 2.5, which gives a lower bound on the probability that the random walk V makes an n-step excursion away from the origin conditioned on the large deviation event $\{Y_n = qn, V_n = 0\}$. Lemmas 4.1 and 4.4 are proven in Section 4.3 whereas the proof of Proposition 2.5 is postponed to Section 6.2 because it requires more preparation. Section 4.4 is dedicated to the proof of Theorem E and Section 4.5 to the proof of Theorem F.

4.1. *Proof of Theorem* C (*One bead theorem*). The proof of Theorem C will be displayed subject to Lemma 4.1 and Proposition 4.2 that are stated below.

LEMMA 4.1. *For $\beta > \beta_c$, there exist $a, a_1, a_2 > 0$ such that*

$$P_{L,\beta}\big(N_L(l) \geq a_1\sqrt{L}\big) \leq a_2 e^{-a\sqrt{L}}, \qquad L \in \mathbb{N}. \tag{4.1}$$

Recall (2.6)–(2.8).

PROPOSITION 4.2. *For $\beta > \beta_c$, there exist $c, c_1, c_2 > 0$ and $\kappa > 1/2$ such that*

$$\frac{c_1}{L^\kappa} e^{\beta L - c\sqrt{L}} \leq Z^\circ_{L,\beta} \leq \frac{c_2}{\sqrt{L}} e^{\beta L - c\sqrt{L}}, \qquad L \in \mathbb{N}. \tag{4.2}$$

4.1.1. *Proof of Theorem* C. We will first show that, for $\beta > \beta_c$ and under the polymer measure, the probability that there is exactly one macroscopic bead in the polymer tends to 1 as $L \to \infty$. Then we will show that, with a probability converging to 1 as $L \to \infty$, the first step and the last step of this macroscopic bead are at distance less than $(\log L)^4$ from 0 and L, respectively. For $r \in \mathbb{N}$, we denote by $Z_{L,\beta}[r]$ the partition function restricted to those trajectories that do not have any bead larger than r, that is,

$$Z_{L,\beta}[r] = \sum_{l \in \Omega_L : |I_{j\max}| \le r} e^{\beta H_L(l)}. \tag{4.3}$$

At this stage, we pick $s > 0$ and we let $\mathcal{A}_{L,s}$ be the subset consisting of those trajectories having at most one bead larger than $s(\log L)^2$, that is,

$$\mathcal{A}_{L,s} = \{l \in \Omega_L : |\{j \in \{1, \dots, n_L(l)\} : |I_j| \ge s(\log L)^2\}| \le 1\}. \tag{4.4}$$

Partition $\mathcal{A}^c_{L,s}$ with respect to the locations of the two subintervals $\{i_1 + 1, \dots, i_2\}$ and $\{i_3 + 1, \dots, i_4\}$ associated with the first two beads that are larger than $s(\log L)^2$. For notational convenience we let $L_1 := i_2 - i_1$ and $L_2 := i_4 - i_3$ be the length of these two first large beads. We do not have Markov property but, with the help of Lemma 4.3 below, we can estimate the partition function restricted to those trajectory that make a bead between two given steps.

Recall (cf. notation introduced in Section 1.3 prior to Theorem C) that x_1 denotes the horizontal extension of the first bead, and that u_{x_1} corresponds to its total length.

LEMMA 4.3. *For* $L \in \mathbb{N}$,

$$\begin{aligned}\tfrac{1}{2} Z^{\circ}_{L',\beta} Z_{L-L',\beta} &\le Z_{L,\beta}(u_{x_1} = L') \\ &\le Z^{\circ}_{L',\beta} Z_{L-L',\beta} \qquad \text{for } L' \in \{1, \dots, L\}.\end{aligned} \tag{4.5}$$

PROOF. In the case $u_{x_1} = 1$, the first bead contains only one horizontal step, hence the sign of the stretch after x_1 is arbitrary, so that obviously $Z_{L,\beta}(u_{x_1} = 1) = Z^{\circ}_{1,\beta} Z_{L-1,\beta}$. In case $u_{x_1} = L' > 1$, note that the stretch l_{x_1} is nonzero, therefore the next stretch has the same sign as l_{x_1}. By concatenating the trajectories,

$$\begin{aligned} &Z_{L,\beta}(u_{x_1} = L') \\ &\quad = Z^{\circ}_{L',\beta}(l_{N_{L'}} > 0) Z_{L-L',\beta}(l_1 \ge 0) + Z^{\circ}_{L',\beta}(l_{N_{L'}} < 0) Z_{L-L',\beta}(l_1 \le 0) \\ &\quad = Z^{\circ}_{L',\beta} Z_{L-L',\beta}(l_1 \ge 0).\end{aligned} \tag{4.6}$$

In both cases, thanks to the symmetry of the stretches, we have

$$\begin{aligned}\tfrac{1}{2} Z^{\circ}_{L',\beta} Z_{L-L',\beta} &\le Z_{L,\beta}(u_{x_1} = L') \\ &\le Z^{\circ}_{L',\beta} Z_{L-L',\beta} \qquad \text{for } L' \in \{1, \dots, L\}.\end{aligned} \tag{4.7}$$

□

We resume the proof of Theorem C and, we use Lemma 4.3 to obtain

$$
\begin{aligned}
&P_{L,\beta}(\mathcal{A}^c_{L,s}) \\
&\quad \le \sum_{\substack{1\le i_1<i_2<i_3<i_4\le L\\ L_1,L_2\ge s(\log L)^2}} \frac{Z_{i_1,\beta}[s(\log L)^2]Z^\circ_{L_1,\beta}Z_{i_3-i_2,\beta}[s(\log L)^2]Z^\circ_{L_2,\beta}Z_{L-i_4,\beta}}{Z_{L,\beta}},
\end{aligned} \tag{4.8}
$$

and we write the lower bound

$$
Z_{L,\beta} \ge \left(\tfrac{1}{2}\right)^3 Z_{i_1,\beta}[s(\log L)^2]Z^\circ_{L_1+L_2,\beta}Z_{i_3-i_2,\beta}[s(\log L)^2]Z_{L-i_4,\beta} \tag{4.9}
$$

such that

$$
P_{L,\beta}(\mathcal{A}^c_{L,s}) \le 8 \sum_{\substack{1\le i_1<i_2<i_3<i_4\le L\\ L_1,L_2\ge s(\log L)^2}} \frac{Z^\circ_{L_1,\beta}Z^\circ_{L_2,\beta}}{Z^\circ_{L_1+L_2,\beta}}. \tag{4.10}
$$

By using Proposition 4.2 and the convex inequality

$$
\sqrt{L_1}+\sqrt{L_2}-\sqrt{L_1+L_2} \ge \tfrac{1}{2}\sqrt{\min\{L_1,L_2\}}, \tag{4.11}
$$

we can bound from above the quantity in the sum in (4.10) by

$$
\frac{Z^\circ_{L_1,\beta}Z^\circ_{L_2,\beta}}{Z^\circ_{L_1+L_2,\beta}} \le \frac{c_1^2(L_1+L_2)^\kappa}{c_2\sqrt{L_1L_2}} e^{-\widetilde{G}(a_\beta)[\sqrt{L_1}+\sqrt{L_2}-\sqrt{L_1+L_2}]} \tag{4.12}
$$

$$
\le \frac{c_1^2(L_1+L_2)^\kappa}{c_2\sqrt{L_1L_2}} e^{-\widetilde{G}(a_\beta)\sqrt{s}\log L/2} \tag{4.13}
$$

and since $\frac{(L_1+L_2)^\kappa}{\sqrt{L_1L_2}} \le L^\kappa$ we can state that, for L large enough, (4.10) becomes

$$
P_{L,\beta}(\mathcal{A}^c_{L,s}) \le \frac{8c_1^2}{c_2} L^{\kappa+4} e^{-\widetilde{G}(a_\beta)\sqrt{s}\log L/2}. \tag{4.14}
$$

Therefore, it suffices to choose $\sqrt{s} = \frac{4(\kappa+4)}{c}$ to conclude that

$$
\lim_{L\to\infty} P_{L,\beta}(\mathcal{A}^c_{L,s}) = 0.
$$

At this stage, we set $\mathcal{B}_{L,s} = \mathcal{A}_{L,s} \cap \{N_L(l) \le a_1\sqrt{L}\}$ and we can use Lemma 4.1 and the fact that $P_{L,\beta}(\mathcal{A}^c_{L,s})$ vanishes as $L\to\infty$ to conclude that $\lim_{L\to\infty} P_{L,\beta}(\mathcal{B}_{L,s}) = 1$. Moreover, it comes easily that under the event $\mathcal{B}_{L,s}$ there is exactly one bead larger than $s(\log L)^2$ because if there were no bead larger than $s(\log L)^2$, then the total number of beads $n_L(l)$ would be larger than $\frac{L}{s(\log L)^2}$ which contradicts the fact that $N_L(l) \le a_1\sqrt{L}$ because each bead contains at least one horizontal step and consequently $N_L(l) \ge n_L(l)$. Under the event $\mathcal{B}_{L,s}$ we denote by i_1 and i_2 the end-steps of the maximal bead, that is, $I_{j_{\max}} = \{i_1+1,\dots,i_2\}$.

Then the proof of Theorem C will be complete once we show that there exists a $v > 0$ such that

$$\lim_{L\to\infty} P_{L,\beta}\big(\mathcal{B}_{L,s} \cap \{i_1 \geq v(\log L)^4\}\big) = 0, \tag{4.15}$$

$$\lim_{L\to\infty} P_{L,\beta}\big(\mathcal{B}_{L,s} \cap \{i_2 \leq L - v(\log L)^4\}\big) = 0. \tag{4.16}$$

We can bound from above

$$\begin{aligned}
&P_{L,\beta}\big(\mathcal{B}_{L,s} \cap \{i_1 \geq v(\log L)^4\}\big) \\
&\quad = \sum_{t=v(\log L)^4}^{L} P_{L,\beta}\big(\mathcal{B}_{L,s} \cap \{i_1 = t\}\big) \\
&\quad \leq \sum_{t=v(\log L)^4}^{L} P_{L,\beta}\big(\exists j \in \{1,\dots,n_L(l)\}: u_{x_j} = t, \\
&\qquad |I_d| \leq s(\log L)^2\ \forall d \in \{1,\dots,j\}\big) \\
&\quad \leq \frac{1}{2} \sum_{t=v(\log L)^4}^{L} \frac{Z_{t,\beta}[s(\log L)^2] Z_{L-t,\beta}}{Z_{t,\beta} Z_{L-t,\beta}},
\end{aligned}$$

which finally gives

$$P_{L,\beta}\big(B_{L,s} \cap \{i_1 \geq v(\log L)^4\}\big) \leq \frac{1}{2} \sum_{t=v(\log L)^4}^{L} P_{t,\beta}\big(|I_{j_{\max}}| \leq s(\log L)^2\big). \tag{4.17}$$

We note that, under $P_{t,\beta}$ and on the event $\{|I_{j_{\max}}| \leq s(\log L)^2\}$, the number of beads is larger than $\frac{t}{s(\log L)^2}$, therefore, $N_t(l) \geq \frac{t}{s(\log L)^2}$ and since $\sqrt{t} \geq \sqrt{v}(\log L)^2$ we obtain that $N_t(l) \geq \sqrt{t}(\sqrt{v}/s)$. By choosing $v = (a_1 s)^2$, we can apply Lemma 4.1 to get

$$\begin{aligned}
P_{L,\beta}\big(\mathcal{B}_{L,s} \cap \{i_1 \geq v(\log L)^4\}\big) &\leq \frac{1}{2} \sum_{t=v(\log L)^4}^{L} P_{t,\beta}\big(N_t(l) \geq a_1\sqrt{t}\big) \\
&\leq \frac{1}{2} a_2 \sum_{t=v(\log L)^4}^{L} e^{-a\sqrt{t}}.
\end{aligned} \tag{4.18}$$

Since the sum in (4.18) vanishes as $L \to \infty$, the proof is complete.

4.2. *Proof of Proposition* 4.2. We recall the definition of the one bead partition function introduced in Section 2.1, equations (2.5)–(2.8). Henceforth, we will use the notation $\widetilde{Z}^{\circ}_{L,\beta} = Z^{\mathsf{m},\circ}_{L,\beta} e^{-\beta L}/c_\beta$, so that Proposition 4.2 will be proven once we show that there exist $c_1, c_2 > 0$ and $\kappa > 1/2$ such that

$$\frac{c_1}{L^\kappa} e^{-\widetilde{G}(a_\beta)\sqrt{L}} \le \widetilde{Z}^{\circ}_{L,\beta} \le \frac{c_2}{\sqrt{L}} e^{-\widetilde{G}(a_\beta)\sqrt{L}} \qquad \text{for } L \in \mathbb{N}\,. \tag{4.19}$$

We will prove (4.19) subject to Lemma 4.4 below and Proposition 2.5. The proof of Lemma 4.4 is given in Section 4.3 whereas the proof of Proposition 2.5 is postponed to Section 6.2. For $K \subset \{1, \dots, L\}$, we set

$$\widetilde{Z}^{\circ}_{L,\beta}(N \in K) = 2 \sum_{N \in K} \big(\Gamma(\beta)\big)^N \mathbf{P}_\beta\big(\mathcal{V}^+_{N+1,L-N}\big), \tag{4.20}$$

and similarly we have

$$\widetilde{Z}^{\circ}_{L,\beta} = 2 \sum_{N=1}^{L} \big(\Gamma(\beta)\big)^N \mathbf{P}_\beta\big(\mathcal{V}^+_{N+1,L-N}\big). \tag{4.21}$$

LEMMA 4.4. *For $\beta > \beta_c$, there exists $a_2 > a_1 > 0$ such that for $L \in \mathbb{N}$,*

$$\lim_{L\to\infty} \frac{\widetilde{Z}^{\circ}_{L,\beta}(a_1\sqrt{L} \le N \le a_2\sqrt{L})}{\widetilde{Z}^{\circ}_{L,\beta}} = 1. \tag{4.22}$$

By using Lemma 4.4, we note that it suffices to prove (4.19) with $\widetilde{Z}^{\circ}_{L,\beta}(N \in \sqrt{L}[a_1, a_2])$ instead of $\widetilde{Z}^{\circ}_{L,\beta}$. For the ease of notation, we will rather take a_2 a bit larger and consider $\widetilde{Z}^{\circ}_{L,\beta}(1 + N \in \sqrt{L}[a_1, a_2])$. In view of (4.20), we write

$$\begin{aligned} &\widetilde{Z}^{\circ}_{L,\beta}\big(1 + N \in \sqrt{L}[a_1, a_2]\big) \\ &\qquad = 2 \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} \big(\Gamma(\beta)\big)^{N-1} \mathbf{P}_\beta\big(\mathcal{V}^+_{N,L-N+1}\big). \end{aligned} \tag{4.23}$$

For $n \in \mathbb{N}$, we recall (1.37) and (2.21) and we note that $nY_n = A_n$ on the set $\{V_n = 0, V_i > 0\ \forall i \in [1, N-1] \cap \mathbb{N}\}$. Therefore, we set $q_{N,L} := \frac{L-N+1}{N^2}$ for $N \in \sqrt{L}[a_1, a_2] \cap \mathbb{N}$ and we can write

$$\mathcal{V}^+_{N,L-N+1} = \big\{V : Y_N = N q_{N,L},\, V_N = 0,\, V_i > 0\ \forall i \in [1, N-1] \cap \mathbb{N}\big\}. \tag{4.24}$$

At this stage, our aim is to bound from above and below the quantities $\mathbf{P}_\beta(\mathcal{V}^+_{N,L-N+1})$ for $N \in \sqrt{L}[a_1, a_2] \cap \mathbb{N}$. The upper bound is obvious, that is,

$$\mathbf{P}_\beta\big(\mathcal{V}^+_{N,L-N+1}\big) \le \mathbf{P}_\beta(Y_N = N q_{N,L}, V_N = 0), \tag{4.25}$$

while the lower bound is obtained as follows. Since $q_{N,L} \in [\frac{1}{2a_2^2}, \frac{1}{a_1^2}]$ when $N \in \sqrt{L}[a_1, a_2]$, we can apply Proposition 2.5 to claim that, there exists $C, \mu > 0$ such that for L large enough,

$$\mathbf{P}_\beta(\mathcal{V}^+_{N,L-N+1}) \geq \frac{C}{N^\mu}\mathbf{P}_\beta(Y_N = Nq_{N,L}, V_N = 0), \qquad N \in \sqrt{L}[a_1, a_2] \cap \mathbb{N}. \tag{4.26}$$

By using again the fact that $q_{N,L} \in [\frac{1}{2a_2^2}, \frac{1}{a_1^2}]$ when $N \in \sqrt{L}[a_1, a_2]$, we can apply Proposition 2.4, which provides a lower and an upper bound on $\mathbf{P}_\beta(Y_N = Nq_{N,L}, V_N = 0)$. By combining these last two bounds with (4.25)–(4.26) and by setting $\kappa = 1 + \mu/2$, we can assert that there exists $R_1 > R_2 > 0$ such that for L large enough and all $N \in \sqrt{L}[a_1, a_2]$ we have that

$$\frac{R_2}{L^\kappa} e^{N[-\widetilde{h}_0(q_{N,L},0)q_{N,L} + \mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q_{N,L},0))]} \leq \mathbf{P}_\beta(\mathcal{V}^+_{N,L-N+1}) \leq \frac{R_1}{L} e^{N[-\widetilde{h}_0(q_{N,L},0)q_{N,L} + \mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q_{N,L},0))]}. \tag{4.27}$$

At this stage, we recall the definition of $\widetilde{G}$ in (1.27) and we set

$$Q_{L,\beta} := \sum_{N=a_1\sqrt{L}}^{a_2\sqrt{L}} e^{\sqrt{L} G_{L,N}} \tag{4.28}$$

with

$$G_{L,N} = \frac{N}{\sqrt{L}}(q_{N,L})^{1/2}\widetilde{G}\left(\frac{1}{(q_{N,L})^{1/2}}\right) \tag{4.29}$$

and we use (4.20) and (4.27) to claim that there exists $R_3 > R_4 > 0$ (depending on β only) such that for L large enough,

$$\frac{R_4}{L^\kappa} Q_{L,\beta} \leq \widetilde{Z}^\circ_{L,\beta}(N \in \sqrt{L}[a_1, a_2]) \leq \frac{R_3}{L} Q_{L,\beta}. \tag{4.30}$$

We recall that $a \mapsto \widetilde{G}(a)$ is a strictly negative and strictly concave function on $(0, \infty)$ and reaches its unique maximum at a_β, which obviously belongs to $[a_1, a_2]$. Since, by Lemma 5.3, $a \mapsto \widetilde{G}(a)$ is $\mathcal{C}^1$ on $(0, \infty)$, we can assert that it is Lipschitz on each compact subset of $(0, \infty)$. Moreover, there exists a $C > 0$ such that $|q_{N+1,L} - q_{N,L}| \leq C/\sqrt{L}$ for $N \in \sqrt{L}[a_1, a_2]$ and we have that

$$\left(1 - \frac{a_2}{\sqrt{L}}\right)^{1/2} \leq \frac{N}{\sqrt{L}}(q_{N,L})^{1/2} \leq \left(1 - \frac{a_1}{\sqrt{L}}\right)^{1/2}, \qquad N \in \sqrt{L}[a_1, a_2], \tag{4.31}$$

therefore, we can take the supremum of $G_{L,N}$ on $N \in [a_1\sqrt{L}, a_2\sqrt{L}] \cap \mathbb{N}$ and it comes that

$$\sup\{G_{L,N}; N \in \sqrt{L}[a_1, a_2] \cap \mathbb{N}\} = \widetilde{G}(a_\beta) + O\left(\frac{1}{\sqrt{L}}\right). \tag{4.32}$$

By putting together (4.28) and (4.32), we obtain that there exists $R_5 > R_6 > 0$ such that for L large enough,

$$R_6 e^{\widetilde{G}(a_\beta)\sqrt{L}} \le Q_{L,\beta} \le R_5 \sqrt{L} e^{\widetilde{G}(a_\beta)\sqrt{L}}. \tag{4.33}$$

At this stage, it suffices to combine (4.30) with (4.33) to complete the proof of (4.19) with $\kappa = \mu/2 + 1$.

4.3. *Proof of Lemmas 4.1 and 4.4.* We will only display the proof of Lemma 4.4 because the proof of Lemma 4.1 is obtained in a very similar manner. We recall (4.20) and (4.21) and we will first show that there exists $\gamma > 0$ and $c > 0$ such that

$$\widetilde{Z}^{\circ}_{L,\beta} \ge c e^{-\gamma\sqrt{L}}, \qquad L \in \mathbb{N}. \tag{4.34}$$

Then we will show that there exist $a_2 > a_1 > 0$ and $c_1, c_2 > 0$ such that

$$\begin{aligned} \widetilde{Z}^{\circ}_{L,\beta}(N \ge a_2\sqrt{L}) &\le c_2 e^{-2\gamma\sqrt{L}}, \qquad L \in \mathbb{N}, \\ \widetilde{Z}^{\circ}_{L,\beta}(N \le a_1\sqrt{L}) &\le c_1 e^{-2\gamma\sqrt{L}}, \qquad L \in \mathbb{N}. \end{aligned} \tag{4.35}$$

Putting together (4.34) and (4.35), we will immediately obtain (4.22). To begin with, set $r := \lfloor \frac{L}{1+\lfloor\sqrt{L}\rfloor} \rfloor$, $u := L - r - (r-1)\lfloor\sqrt{L}\rfloor$ and note that $u \in \{\lfloor\sqrt{L}\rfloor, \dots, 2\lfloor\sqrt{L}\rfloor\}$. Then consider the trajectory $V^* \in \mathcal{V}^+_{r+1,L-r}$ defined as $V_0 = V_{r+1} = 0$, $V_1 = \cdots = V_{r-1} = \lfloor\sqrt{L}\rfloor$ and $V_r = u$. One can therefore compute

$$\mathbf{P}_\beta(V^*) = \left(\frac{1}{c_\beta}\right)^{r+1} e^{-(\beta/2)(2u)} \ge \left(\frac{1}{c_\beta}\right)^{r+1} e^{-2\beta\lfloor\sqrt{L}\rfloor}, \tag{4.36}$$

and consequently by restricting the sum in (4.20) to $N = r$, by using (4.36) and the inequality $\lfloor\sqrt{L}\rfloor \le \sqrt{L}$, we obtain

$$\widetilde{Z}^{\circ}_{L,\beta} \ge \frac{2}{c_\beta}\left(\frac{\Gamma(\beta)}{c_\beta}\right)^{r} e^{-2\beta\sqrt{L}}. \tag{4.37}$$

It remains to note that $r \le \sqrt{L}$ and to recall that $c_\beta > 1$ and that $\Gamma(\beta) < 1$ because $\beta > \beta_c$. This is sufficient to obtain (4.34).

Proving the first inequality in (4.35) is easy because $\Gamma(\beta) < 1$, and thus, we can use (4.20) to claim that there exists a $C > 0$ such that

$$\widetilde{Z}^{\circ}_{L,\beta}(N \ge a_2\sqrt{L}) \le 2 \sum_{N=a_2\sqrt{L}}^{\infty} \big(\Gamma(\beta)\big)^N \le C e^{a_2 \log(\Gamma(\beta))\sqrt{L}}. \tag{4.38}$$

Since $\log(\Gamma(\beta)) < 0$, it suffices to choose a_2 large enough to obtain the first inequality in (4.35).

To prove the last inequality in (4.35), we note that, for $N \le a_1\sqrt{L}$ and for all $(V_i)_{i=0}^{N+1} \in \mathcal{V}^+_{N+1,L-N}$ we have $\max\{V_j, j \in \{1,\dots,N\}\} \ge \frac{L-N}{N} \ge \frac{\sqrt{L}}{a_1} - 1$ and therefore, for L large enough we have

$$\mathbf{P}_\beta(\mathcal{V}^+_{N+1,L-N}) \le \mathbf{P}_\beta\left(\max\{V_j, j \le a_1\sqrt{L}\} \ge \frac{\sqrt{L}}{2a_1}\right) \tag{4.39}$$

$$\le \mathbf{P}_\beta\left(\sum_{i=1}^{a_1\sqrt{L}} |U_i| > \frac{\sqrt{L}}{2a_1}\right), \tag{4.40}$$

and since U_1 has some finite exponential moments, we can apply a standard Cramér's theorem to obtain that for L large enough, there exists $g(a_1) > 0$ such that $\lim_{a_1\to 0^+} g(a_1) = \infty$ and that $\mathbf{P}_\beta(\mathcal{V}^+_{N+1,L-N}) \le e^{-g(a_1)\sqrt{L}}$ for $N \le a_1\sqrt{L}$. Therefore, by taking a_1 small enough we obtain the second inequality in (4.35), which completes the proof of Lemma 4.4.

4.4. *Proof of Theorem* E (*Horizontal extension*). To begin this section, we prove that $\widetilde{G}$ is strictly concave and reaches its maximum at a unique point $a_\beta \in (0,\infty)$. Recall (1.27) and compute its first two derivatives (by using that $\nabla\mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q,0)) = (q,0)$), that is,

$$\frac{d}{da}\widetilde{G}(a) = \log\Gamma(\beta) + \frac{1}{a^2}\widetilde{h}_0\left(\frac{1}{a^2},0\right) + \mathfrak{L}_\Lambda\left(\widetilde{\mathbf{h}}\left(\frac{1}{a^2},0\right)\right), \tag{4.41}$$

$$\frac{d^2}{da^2}\widetilde{G}(a) = -\frac{2}{a^3}\widetilde{h}_0\left(\frac{1}{a^2},0\right) - \frac{4}{a^5}\partial_1\widetilde{h}_0\left(\frac{1}{a^2},0\right). \tag{4.42}$$

It suffices to show that $\frac{d^2}{da^2}\widetilde{G}(a) < 0$ on $(0,\infty)$ and that $\frac{d}{da}\widetilde{G}(a)$ has a zero on $(0,\infty)$. Since $\widetilde{h}_0(x,0) = -2\widetilde{h}_1(x,0)$ (recall Remark 5.5), we consider $R : u \mapsto \int_0^1 x\mathfrak{L}'((x-\frac{1}{2})u)\,dx$ so that $\partial_1(\mathfrak{L}_\Lambda)(\widetilde{\mathbf{h}}(x,0)) = R(\widetilde{h}_0(x,0))$. Clearly, $R(0) = 0$ and $R'(u) = 2\int_0^1 x^2\mathfrak{L}''(xu)\,dx$ because $\mathfrak{L}$ is even [recall (1.23)]. Therefore, $R'(u) > 0$ when $u \ne 0$ and $R < 0$ on $(-\infty,0)$ and $R > 0$ on $(0,\infty)$. Since $R(\widetilde{h}_0(x,0)) = x$ for $x \in \mathbb{R}$, we can claim that $\widetilde{h}_0(x,0) > 0$ for $x \in (0,\infty)$ and by differentiating this latter equality we obtain that $\partial_1\widetilde{h}_0(x,0) = 1/R'(\widetilde{h}_0(x,0))$, which is strictly positive on $(0,\infty)$. This completes the proof.

Let us start the proof of Theorem E. Recall that i_1 and i_2 are the end-steps of the largest bead $I_{j_{\max}}$, that is, $I_{j_{\max}} = \{i_1+1,\dots,i_2\}$. For $v > 0$, we let

$$\begin{aligned}\mathcal{T}_{L,v} := \{&l \in \Omega_L : i_1 \le v(\log L)^4, i_2 \ge L - v(\log L)^4,\\ &I_{j_{\max}} = \{i_1+1,\dots,i_2\}\}.\end{aligned} \tag{4.43}$$

By Theorem C, there exists a $v > 0$ such that $\lim_{L\to\infty} P_{L,\beta}(\mathcal{T}_{L,v}) = 1$. Therefore, the proof will be complete once we show that

$$\lim_{L\to\infty} P_{L,\beta}\left(\left\{\left|\frac{N_L(l)}{\sqrt{L}} - a_\beta\right| > \varepsilon\right\} \cap \mathcal{T}_{L,v}\right) = 0. \tag{4.44}$$

Let $N_{I_{j_{\max}}}$ denote the number of horizontal steps made by the random walk in its largest bead. Pick $\varepsilon' < \varepsilon$ and since the first step and the last step of the largest bead are at distance less than $v(\log L)^4$ from 0 and L, respectively, we can write that for L large enough

$$P_{L,\beta}\Big(\Big\{\Big|\frac{N_L(l)}{\sqrt{L}} - a_\beta\Big| > \varepsilon\Big\} \cap \mathcal{T}_{L,v}\Big)$$

$$\leq \sum_{\substack{1 \leq i_1 \leq v(\log L)^4 \\ L - v(\log L)^4 \leq i_2 \leq L}} P_{L,\beta}\Big(\Big|\frac{N_{I_{j_{\max}}}}{\sqrt{i_2 - i_1}} - a_\beta\Big| > \varepsilon', I_{j_{\max}} = \{i_1 + 1, \dots, i_2\}\Big) \tag{4.45}$$

$$\leq 4 \sum_{\substack{1 \leq i_1 \leq v(\log L)^4 \\ L - v(\log L)^4 \leq i_2 \leq L}} \frac{Z^{\circ}_{i_2 - i_1, \beta}(|(N/\sqrt{i_2 - i_1}) - a_\beta| > \varepsilon')}{Z^{\circ}_{i_2 - i_1, \beta}},$$

where the coefficient 4 in front of the RHS in (4.45) comes from a direct application of Lemma 4.3. Now, we focus on the numerator of the RHS in (4.45) and since $\widetilde{G}$ is strictly concave and reaches its maximum at a_β we can claim that the maximum of $\widetilde{G}$ on $(0, a_\beta - \varepsilon'] \cup [a_\beta + \varepsilon', \infty)$ is given by $T(\varepsilon') = \max\{\widetilde{G}(a_\beta - \varepsilon'), \widetilde{G}(a_\beta + \varepsilon')\}$. We proceed as in (4.23)–(4.32) and we get that there exist a $C_1 > 0$ such that

$$Z^{\circ}_{i_2 - i_1, \beta}\Big(\Big|\frac{N}{\sqrt{i_2 - i_1}} - a_\beta\Big| > \varepsilon'\Big) \leq \frac{C_1}{\sqrt{i_2 - i_1}} e^{\beta(i_2 - i_1)} e^{T(\varepsilon')\sqrt{i_2 - i_1}}. \tag{4.46}$$

We apply Proposition 4.2 and the denominator can be bounded from below as

$$Z^{\circ}_{i_2 - i_1, \beta} \geq \frac{C_2}{(i_2 - i_1)^{\kappa}} e^{\beta(i_2 - i_1)} e^{\widetilde{G}(a_\beta)\sqrt{i_2 - i_1}}, \tag{4.47}$$

for some constants $\kappa > 1/2$ and $C_2 > 0$. Since $L - 2v(\log L)^4 \leq i_2 - i_1 \leq L$, we can state that, for L large enough, (4.45) becomes

$$\begin{aligned} &P_{L,\beta}\Big(\Big\{\Big|\frac{N_L(l)}{\sqrt{L}} - a_\beta\Big| > \varepsilon\Big\} \cap \mathcal{T}_{L,v}\Big) \\ &\quad \leq C_3 L^{\kappa - 1/2} \log^8(L) e^{-(\widetilde{G}(a_\beta) - T(\varepsilon'))\sqrt{L - 2v \log^4 L}}. \end{aligned} \tag{4.48}$$

Since $\widetilde{G}(a_\beta) > T(\varepsilon')$, the RHS vanishes as $L \to \infty$, and this completes the proof.

4.5. *Proof of Theorem* F (*Wulff shape*). Before displaying the proof of Theorem F, we provide a rigorous definition of γ^*_β and we associate with each trajectory $l \in \Omega_L$ the process M_l that links the middle of each stretch consecutively.

The Wulff shape γ_β^* can be defined[3] as

$$\gamma_\beta^* = \operatorname{argmin}\left\{ J(\gamma), \gamma \in \mathcal{B}_{[0,1]}, \int_0^1 \gamma(t)\,dt = \frac{1}{a_\beta^2}, \gamma(0) = \gamma(1) = 0 \right\}, \tag{4.49}$$

where $\mathcal{B}_{[0,1]}$ is the set containing the cadlag real functions defined on [0, 1], where $J : \mathcal{B}_{[0,1]} \to [0, \infty]$ is defined as

$$J(\gamma) = \begin{cases} \int_0^1 \mathfrak{L}^*(\gamma'(t))\,dt, & \text{if } \gamma \in \mathcal{AC}, \\ +\infty, & \text{otherwise}, \end{cases} \tag{4.50}$$

where $\mathcal{AC}$ is the set of absolutely continuous functions and where $\mathfrak{L}^*$ is the Legendre transform of $\mathfrak{L}$, that is,

$$\mathfrak{L}^*(u) = \sup\left\{ hu - \mathfrak{L}(h), h \in \left(-\frac{\beta}{2}, \frac{\beta}{2}\right) \right\}, \qquad u \in \mathbb{R}. \tag{4.51}$$

Using the duality between $\mathfrak{L}$ and $\mathfrak{L}^*$, we easily obtain the formula (1.28) given in the Introduction, which easily implies [recall (1.27)] that $\widetilde{G}(a_\beta) = a_\beta(\log \Gamma(\beta) - J(\gamma_\beta^*))$. Finally, we note that one can prove without further difficulty that

$$\{-\gamma_\beta^*, \gamma_\beta^*\} = \operatorname{argmin}\left\{ J(\gamma), \gamma \in \mathcal{B}_{[0,1]}, A(\gamma) = \frac{1}{a_\beta^2}, \gamma(0) = \gamma(1) = 0 \right\}, \tag{4.52}$$

where $A(\gamma) := \int_0^1 |\gamma(s)|\,ds$ is the geometric area enclosed between the graph of γ and the x-axis.

We recall the definition of $\mathcal{E}_l^+$ and $\mathcal{E}_l^-$ in (1.32) and we also associate with each $l \in \mathcal{L}_{N,L}$ the path $M_l = (M_{l,i})_{i=0}^{N+1}$ that links the middles of each stretch consecutively and is defined as $M_{l,0} = 0$

$$M_{l,i} = l_1 + \cdots + l_{i-1} + \frac{l_i}{2}, \qquad i \in \{1, \ldots, N\}, \tag{4.53}$$

and $M_{l,N+1} = l_1 + \cdots + l_N$. We recall that the T_N transformation, defined in Section 2.1, associates with each $l \in \mathcal{L}_{N,L}$ the path $V_l = (T_N)^{-1}(l)$ such that $V_{l,0} = 0$, $V_{l,i} = (-1)^{i-1} l_i$ for all $i \in \{1, \ldots, N\}$ and $V_{l,N+1} = 0$. As a consequence, $\mathcal{E}_l^+ = M_l + \frac{|V_l|}{2}$ and $\mathcal{E}_l^- = M_l - \frac{|V_l|}{2}$, that is,

$$\begin{aligned} \mathcal{E}_{l,i}^+ &= M_{l,i} + \frac{|V_{l,i}|}{2}, \qquad i \in \{0, \ldots, N+1\}, \\ \mathcal{E}_{l,i}^- &= M_{l,i} - \frac{|V_{l,i}|}{2}, \qquad i \in \{0, \ldots, N+1\}, \end{aligned} \tag{4.54}$$

[3]The set on the RHS of (4.49) is not empty since it contains the hat function $\gamma(t) = \gamma(1-t) = \frac{2t}{a_\beta}$ for $0 \le t \le \frac{1}{2}$.

and the path $(M_{l,i})_{i=0}^{N+1}$ can be rewritten with the increments $(U_i)_{i=1}^{N+1}$ of the V_l random walk as

$$M_{l,i} = \sum_{j=1}^{i} (-1)^{j+1} \frac{U_j}{2}, \qquad i \in \{1, \dots, N\}. \tag{4.55}$$

Similarly to what we did to define $\widetilde{\mathcal{E}}_l^+$ and $\widetilde{\mathcal{E}}_l^-$ in (1.34), we let $\widetilde{M}_l$ and $\widetilde{V}_l$ be the time–space rescaled cadlag process associated to M_l and V_l.

PROOF OF THEOREM F. Equations (4.54) that allows to express $\mathcal{E}_l^+$ and $\mathcal{E}_l^-$ with the help of the two processes V_l and M_l can be translated in terms of the time–space rescaled processes as $\widetilde{\mathcal{E}}_l^+ = \widetilde{M}_l + \frac{|\widetilde{V}_l|}{2}$ and $\widetilde{\mathcal{E}}_l^- = \widetilde{M}_l - \frac{|\widetilde{V}_l|}{2}$. Therefore, Theorem F is a straightforward consequence of the two following lemmas.

LEMMA 4.5. *For $\beta > \beta_c$ and $\varepsilon > 0$,*

$$\lim_{L\to\infty} P_{L,\beta}\big(\| |\widetilde{V}_l| - \gamma_\beta^* \|_\infty > \varepsilon\big) = 0. \tag{4.56}$$

LEMMA 4.6. *For $\beta > 0$ and $\varepsilon > 0$,*

$$\lim_{L\to\infty} P_{L,\beta}\big(\|\widetilde{M}_l\|_\infty > \varepsilon\big) = 0. \tag{4.57}$$

□

PROOF OF LEMMA 4.5. For conciseness, we set $\mathcal{U}_{L,\varepsilon} = \{l \in \Omega_L : \| |\widetilde{V}_l| - \gamma_\beta^* \|_\infty > \varepsilon\}$. Thanks to Theorem E, Lemma 4.5 will be proven once we show that there exists an $\eta > 0$ such that

$$\lim_{L\to\infty} P_{L,\beta}\left(\mathcal{U}_{L,\varepsilon} \cap \left\{\left|\frac{N_L(l)}{\sqrt{L}} - a_\beta\right| \le \eta\right\}\right) = 0. \tag{4.58}$$

We decompose the LHS in (4.58) with respect to the value taken by $N_L(l)$, that is,

$$\begin{aligned} &P_{L,\beta}\left(\mathcal{U}_{L,\varepsilon} \cap \left\{\left|\frac{N_L(l)}{\sqrt{L}} - a_\beta\right| \le \eta\right\}\right) \\ &\quad = \sum_{N \in I_{\eta,L}} P_{L,\beta}\big(\mathcal{U}_{L,\varepsilon} \cap \{N_L(l) = N\}\big), \end{aligned} \tag{4.59}$$

where $I_{\eta,L} = \{(a_\beta - \eta)\sqrt{L}, \dots, (a_\beta + \eta)\sqrt{L}\}$. By recalling Section 2.1, the probability in the RHS of (4.59) can be rewritten, with the help of the random walk representation, as

$$\begin{aligned} &P_{L,\beta}\big(\mathcal{U}_{L,\varepsilon} \cap \{N_L(l) = N\}\big) \\ &\quad = \frac{(\Gamma(\beta))^N}{\widetilde{Z}_{L,\beta}} \mathbf{P}_\beta\bigg(\| |\widetilde{V}_{N+1}| - \gamma_\beta^* \|_\infty > \varepsilon, \\ &\qquad \widetilde{V}_{N+1}(1) = 0, A(\widetilde{V}_{N+1}) = \frac{L-N}{(N+1)^2}\bigg), \end{aligned}$$

where $(V_i)_{i=0}^{N+1}$ is a random walk of law $\mathbf{P}_\beta$ and $\widetilde{V}_{N+1}$ is the time–space rescaled process associated with $(V_i)_{i=0}^{N+1}$, that is,

$$\widetilde{V}_{N+1}(t) = \frac{1}{N+1} V_{\lfloor t(N+1) \rfloor}, \qquad t \in [0, 1],$$

and where $\widetilde{Z}_{L,\beta} = Z_{L,\beta} e^{-\beta L}/c_\beta$. Note that there exists a function $g : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\lim_{\eta \to 0} g(\eta) = 0$ and such that for $N \in I_{\eta,L}$ the probability in the RHS of (4.60) is bounded from above by $\mathbf{P}_\beta(\widetilde{V}_N \in \mathcal{H}_{\varepsilon,\eta})$, where

$$\mathcal{H}_{\varepsilon,\eta} = \Big\{ \gamma \in \mathcal{B}_{[0,1]} \colon A(\gamma) \geq \frac{1}{a_\beta^2} - g(\eta), \gamma(0) = \gamma(1) = 0, \\ \| |\gamma| - \gamma_\beta^* \|_\infty \geq \varepsilon \Big\}. \tag{4.60}$$

Thus, we need to identify the exponential growth rate of $\mathbf{P}_\beta(\widetilde{V}_N \in \mathcal{H}_{\varepsilon,\eta})$. To that aim, we apply the Mogulskii theorem [see Dembo and Zeitouni (2010), Theorem 5.1.2] which ensures that $(\widetilde{V}_N)_{N \in \mathbb{N}}$ follows a large deviation principle on the set $\mathcal{B}([0, 1])$ endowed with the supremum norm $\| \cdot \|_\infty$ and with the good rate function J defined in (4.50). Since $\mathcal{H}_{\varepsilon,\eta}$ is a closed subset of $(\mathcal{B}_{[0,1]}, \| \cdot \|_\infty)$ we can assert that

$$\limsup_{n \to \infty} \frac{1}{N} \log \mathbf{P}_\beta(\widetilde{V}_N \in \mathcal{H}_{\varepsilon,\eta}) \leq -\inf\{J(\gamma), \gamma \in \mathcal{H}_{\varepsilon,\eta}\}. \tag{4.61}$$

We pick $M > \inf\{J(\gamma), \gamma \in \mathcal{H}_{\varepsilon,1}\}$ and set $\mathcal{H}_{\varepsilon,\eta}^M = \{\gamma \in \mathcal{H}_{\varepsilon,\eta} \colon J(\gamma) \leq M\}$ such that the inequality (4.61) becomes

$$\limsup_{n \to \infty} \frac{1}{N} \log \mathbf{P}_\beta(\widetilde{V}_N \in \mathcal{H}_{\varepsilon,\eta}) \leq -\inf\{J(\gamma), \gamma \in \mathcal{H}_{\varepsilon,\eta}^M\}. \tag{4.62}$$

At this stage, it remains to show that there exists $\alpha > 0$ and $\eta_0 > 0$ such that for all $\eta \in (0, \eta_0]$,

$$\inf\{J(\gamma), \gamma \in \mathcal{H}_{\varepsilon,\eta}^M\} - \alpha \geq \inf\{J(\gamma), \gamma \in \mathcal{H}_{0,0}\} = J(\gamma_\beta^*). \tag{4.63}$$

Assume that (4.63) fails to be true, then there exists a strictly positive sequence $(z_n)_{n \in \mathbb{N}}$ that tends to 0 as $n \to \infty$ such that for all $n \in \mathbb{N}$ there exists a $\gamma_n \in \mathcal{H}_{\varepsilon,z_n}^M$ satisfying $J(\gamma_n) \leq J(\gamma_\beta^*) + 1/n$. Since J is a good rate function, we can assert that $\mathcal{H}_{\varepsilon,1}^M$ is a compact set of $(\mathcal{B}_{[0,1]}, \| \cdot \|_\infty)$ and consequently γ_n is converging by subsequence toward some $\gamma_\infty \in \mathcal{H}_{\varepsilon,1}^M$. Since A and J are continuous and lower semi-continuous on $(\mathcal{B}_{[0,1]}, \| \cdot \|_\infty)$, respectively, it comes that $\gamma_\infty \in \mathcal{H}_{\varepsilon,0}^M$ and $J(\gamma_\infty) \leq J(\gamma_\beta^*)$, which leads to a contradiction because $-\gamma_\beta^*$ and γ_β^* are the unique maximizer of J on $\mathcal{H}_{0,0}$ and $\gamma_\infty \notin \{-\gamma_\beta^*, \gamma_\beta^*\}$. At this stage, we go back to

(4.60) and we can write, for $\eta \in (0,1]$

$$P_{L,\beta}\big(\mathcal{U}_{L,\varepsilon} \cap \{|N_L(l)-a_\beta| \le \eta\}\big) \le \frac{2\eta}{\widetilde{Z}_{L,\beta}} \sqrt{L}\big(\Gamma(\beta)\big)^{(a_\beta-\eta)\sqrt{L}} \mathbf{P}_\beta(\widetilde{V}_{N+1} \in \mathcal{H}_{\varepsilon,\eta}). \tag{4.64}$$

Thus, by (4.62) and (4.64) we can assert that for all $\eta \in (0,\eta_0]$ and for L large enough

$$\begin{aligned} P_{L,\beta}\big(\mathcal{U}_{L,\varepsilon} \cap \{|N_L(l)-a_\beta| \le \eta\}\big) &\le \frac{2\eta\sqrt{L}}{\widetilde{Z}_{L,\beta}} \big(\Gamma(\beta)\big)^{(a_\beta-\eta)\sqrt{L}} e^{-(a_\beta-\eta)\sqrt{L}(J(\gamma_\beta^*)+\alpha)}, \\ &\le \frac{2\eta\sqrt{L}}{\widetilde{Z}_{L,\beta}} e^{\sqrt{L}(a_\beta-\eta)(\log\Gamma(\beta)-J(\gamma_\beta^*)-\alpha)}. \end{aligned} \tag{4.65}$$

Recall the equality $\widetilde{G}(a_\beta) = a_\beta(\log\Gamma(\beta) - J(\gamma_\beta^*))$ and recall that for $\beta > \beta_c$, we have proved in (4.19) that there exists $c_1 > 0$ and $\kappa > 0$ such that for L large enough,

$$\widetilde{Z}_{L,\beta} \ge \frac{c_1}{L^\kappa} e^{\sqrt{L}\widetilde{G}(a_\beta)}. \tag{4.66}$$

Thus, we can use (4.65) to claim that by choosing η small enough and L large enough we have for a constant $c_2 > 0$,

$$P_{L,\beta}\big(\mathcal{U}_{L,\varepsilon} \cap \{|N_L(l)-a_\beta| \le \eta\}\big) \le \frac{1}{c_2} L^{1/2+\kappa} e^{-(\alpha/2)a_\beta\sqrt{L}}, \tag{4.67}$$

which completes the proof of Lemma 4.5. □

PROOF OF LEMMA 4.6. Lemma 4.6 will be proven once we show that for all $\varepsilon > 0$,

$$\lim_{L\to\infty} P_{L,\beta}\left(\frac{1}{1+N_L(l)} \max_{i \le 1+N_L(l)} |M_{l,i}| \ge \varepsilon\right) = 0. \tag{4.68}$$

Proving (4.68) requires to control, under $P_{L,\beta}$, the probability that, the gap between the modulus of the algebraic area $(N_L(l)|Y_l| := |\sum_{i=1}^{N_L(l)} V_{l,i}|)$ and the geometric area $(\sum_{i=1}^{N_L(l)} |V_{l,i}|)$ of the random walk trajectory $V_l = (T_{N_L(l)})^{-1}(l)$ associated with $l \in \Omega_L$ does not exceed $\log(L)^4$. This is the object of Lemma 4.7 below.

LEMMA 4.7. *For $\beta > \beta_c$ there exists a $c > 0$ such that*

$$\lim_{L\to\infty} P_{L,\beta}\big(N_L(l)|Y_l| \notin [L - N_L(l) - c(\log L)^4, L - N_L(l)]\big) = 0. \tag{4.69}$$

PROOF. By Theorem C, there exists a $c > 0$ such that

$$\lim_{L\to\infty} P_{L,\beta}\big[|I_{j_{\max}}| \leq L - c(\log L)^4\big] = 0. \tag{4.70}$$

Note that for $l \in \Omega_L$, we have $\sum_{i=1}^{N_L(l)} |V_{l,i}| = \sum_{i=1}^{N_L(l)} |l_i| = L - N_L(l)$ and that, with the definition of $j_{\max}$ and $x_{j_{\max}}$ in (1.19) and (1.20) we have also

$$\sum_{i=1}^{N_L(l)} |V_{l,i}| - 2\sum_{i\notin \mathcal{O}_l} |V_{l,i}| \leq \left|\sum_{i=1}^{N_L(l)} V_{l,i}\right| \leq \sum_{i=1}^{N_L(l)} |V_{l,i}|, \tag{4.71}$$

where $\mathcal{O}_l = \{x_{j_{\max}-1} + 1, \dots, x_{j_{\max}}\}$ gathers the indexes of those stretches in $l = (l_1, \dots, l_{N_L(l)})$ that belong to the largest bead described by l. Moreover, we note that $l \in \{|I_{j_{\max}}| \geq L - c(\log L)^4\}$ yields

$$\sum_{i\notin \mathcal{O}_l} |V_{l,i}| = \sum_{i\notin \mathcal{O}_l} |l_i| \leq c(\log L)^4. \tag{4.72}$$

At this stage, we recall that $N_L(l)Y_l = \sum_{i=1}^{N_L(l)} V_{l,i}$ and we use (4.71) and (4.72) to assert that $l \in \{|I_{j_{\max}}| \geq L - c(\log L)^4\}$ implies $N_L(l)|Y_l| \in [L - N_L(l) - 2c(\log L)^4, L - N_L(l)]$. It remains to use (4.70) to complete the proof of Lemma 4.7. □

Let us resume the proof of Lemma 4.6. For $\varepsilon > 0$ and for $\eta > 0$, we set

$$\begin{aligned} K_{L,\varepsilon} &= \left\{\frac{1}{1+N_L(l)} \max_{i\leq 1+N_L(l)} |M_{l,i}| \geq \varepsilon\right\}, \\ R_{L,\eta} &= \left\{\left|\frac{N_L(l)}{\sqrt{L}} - a_\beta\right| \leq \eta\right\} \\ &\quad \cap \big\{N_L(l)|Y_l| \in \big[L - N_L(l) - c(\log L)^4, L - N_L(l)\big]\big\}. \end{aligned} \tag{4.73}$$

Thanks to Theorem E and Lemma 4.7, it suffices to show that there exists $\eta > 0$ such that for all $\varepsilon > 0$,

$$\lim_{L\to\infty} P_{L,\beta}(K_{L,\varepsilon} \cap R_{L,\eta}) = 0. \tag{4.74}$$

We decompose the LHS in (4.74) with respect to the value taken by $N_L(l)$ and Y_l, that is,

$$\begin{aligned} &P_{L,\beta}(K_{L,\varepsilon} \cap R_{L,\eta}) \\ &\quad = \sum_{N\in I_{\eta,L}} \sum_{q\in F_{L,N}} \big[P_{L,\beta}\big(K_{L,\varepsilon} \cap \{N_L(l) = N\} \cap \{Y_l = q(N+1)\}\big) \\ &\qquad + P_{L,\beta}\big(K_{L,\varepsilon} \cap \{N_L(l) = N\} \cap \{Y_l = -q(N+1)\}\big)\big], \end{aligned} \tag{4.75}$$

where

$$I_{\eta,L} = \{(a_\beta - \eta)\sqrt{L}, \dots, (a_\beta + \eta)\sqrt{L}\},$$
$$F_{L,N} = \frac{1}{N(N+1)}\{L - N - c(\log L)^4, \dots, L - N\}.$$

We recall the definition of A_N below (1.14) and of Y_N in (2.21). With the random walk representation we obtain, for $N \in I_{\eta,L}$ and $q \in F_{L,N}$, that

$$\begin{aligned} &P_{L,\beta}\big(K_{L,\varepsilon} \cap \{N_L(l) = N\} \cap \{Y_l = q(1+N)\}\big) \\ &\quad = \frac{(\Gamma(\beta))^N}{\widetilde{Z}_{L,\beta}} \mathbf{P}_\beta\bigg(A_N = L - N, Y_{N+1} = q(N+1), \\ &\qquad \frac{1}{1+N} \max_{i \le 1+N} |M_{N+1,i}| \ge \varepsilon, V_{N+1} = 0\bigg) \\ &\quad \le \frac{(\Gamma(\beta))^N}{\widetilde{Z}_{L,\beta}} \mathbf{P}_\beta\big(Y_{N+1} = q(N+1), V_{N+1} = 0\big) D_{N+1,q}, \end{aligned} \tag{4.76}$$

where $\widetilde{Z}_{L,\beta} = Z_{L,\beta} e^{-\beta L}/c_\beta$, where the middle line $(M_{N+1,i})_{i=0}^{N+1}$ is defined with the increments $(U_i)_{i=1}^{N+1}$ of the V random walk [recall (4.55)] as $M_{N+1,i} = \sum_{j=1}^{i} (-1)^{i+1} \frac{U_i}{2}$ for $i = 1, \dots, N+1$, and where

$$D_{N,q} = \mathbf{P}_\beta\bigg(\frac{1}{N} \max_{i \le N} |M_{N,i}| \ge \varepsilon | Y_N = qN, V_N = 0\bigg). \tag{4.77}$$

By picking $\eta = a_\beta/2$, we can easily check that there exists $[q_1, q_2] \subset (0, \infty)$ such that for all $N \in I_{\eta,L}$ we have $F_{N,L} \subset [q_1, q_2]$. We recall (2.26) and we tilt $\mathbf{P}_\beta$ into $\mathbf{P}_{N,\mathbf{h}_N^q}$ so that we can use Proposition 2.2 and claim that there exists a $c > 0$ such that that for L large enough, we have

$$\begin{aligned} D_{N,q} &\le \frac{\mathbf{P}_{N,\mathbf{h}_N^q}((1/N)\max_{i \le N} |M_{N,i}| \ge \varepsilon)}{\mathbf{P}_{N,\mathbf{h}_N^q}(Y_N = qN, V_{N+1} = 0)} \\ &\le cN^2 \mathbf{P}_{N,\mathbf{h}_N^q}\Big(\max_{i \le N} |M_{N,i}| \ge \varepsilon N\Big). \end{aligned} \tag{4.78}$$

At this stage, we use (4.75), (4.76), (4.78) and the inequalities $\Gamma(\beta) < 1$ and (4.66) to assert that the proof of Lemma 4.6 will be complete once we show that for $[q_1, q_2] \in (0, \infty)$ and $\varepsilon > 0$ there exists a $\vartheta > 0$ such that for N large enough we have

$$\sup_{q \in [q_1, q_2]} \mathbf{P}_{N,\mathbf{h}_N^q}\Big(\max_{i \le N} |M_{N,i}| \ge \varepsilon N\Big) \le e^{-\vartheta N}. \tag{4.79}$$

We recall that, for $1 \le j \le N$, we have $\mathbf{E}_{N,\mathbf{h}_N^q}(U_j) = \mathfrak{L}'(h_N^j)$ with $h_N^j = (1 - \frac{j}{N})h_{N,0}^q + h_{N,1}^q$. As a consequence, and because of Lemma 5.4, we can assert that, for N large enough and uniformly in $q \in [q_1, q_2]$, all h_N^i belong to some compact set $K \subset (-\frac{\beta}{2}, \frac{\beta}{2})$. Therefore, we can show that there exists $c_1 > 0$ and $M_1 > 0$ such that for N large enough

$$\sup_{q \in [q_1, q_2]} \sup_{1 \le i \le N} \mathbf{E}_{N, \mathbf{h}_N^q}\left(e^{c_1 |U_i|}\right) \le M_1, \tag{4.80}$$

which is sufficient to deduce, still for N large enough, that there exists $c_2 > 0$ and $\delta_0 > 0$ such that

$$\sup_{q \in [q_1, q_2]} \sup_{1 \le i \le N} \sup_{\delta \in [-\delta_0, \delta_0]} \mathbf{E}_{N, \mathbf{h}_N^q}\left(e^{\delta (U_i - L'(h_{N,j}))}\right) \le e^{c_2 \delta^2}. \tag{4.81}$$

Then we set

$$\widehat{M}_{N,i} = M_{N,i} - \mathbf{E}_{N, \mathbf{h}_N^q}(M_{N,i}) = \frac{1}{2} \sum_{j=1}^{i} (-1)^{j+1} \big(U_j - \mathfrak{L}'(h_{N,j})\big), \qquad i = 1, \ldots, N, \tag{4.82}$$

and since, under the law $\mathbf{P}_{N, \mathbf{h}_N^q}$, the increments $(U_i)_{i=0}^N$ are independent, we deduce from (4.81) that, for N large enough, there exists $c_3 > 0$ and $\delta_0 > 0$ such that

$$\sup_{q \in [q_1, q_2]} \sup_{1 \le i \le N} \sup_{\delta \in [-\delta_0, \delta_0]} \mathbf{E}_{N, \mathbf{h}_N^q}\left(e^{\delta \widehat{M}_{N,i}}\right) \le e^{c_3 \delta^2 N}. \tag{4.83}$$

The inequality in (4.83) is sufficient to derive (4.79) with random variables $(\widehat{M}_{N,i})_{i=1}^N$ instead of $(M_{N,i})_{i=1}^N$. Then we recover (4.79) by showing that $\mathbf{E}_{N, \mathbf{h}_N^q}(M_{N,i})$ is bounded by some constant uniformly in $q \in [q_1, q_2]$, $N \ge 2$ and $i \in \{1, \ldots, N\}$. The latter boundedness is obtained by writing, for all $1 \le i \le N$ that

$$\begin{aligned} 2\left|\mathbf{E}_{N, \mathbf{h}_N^q}[M_{N,i}]\right| &= \left|\sum_{j=1}^{i} (-1)^j \mathfrak{L}'(h_N^j)\right| \\ &\le \|\mathfrak{L}'\|_{\infty, K} + \left|\sum_{j=1}^{\lfloor i/2 \rfloor} \mathfrak{L}'(h_N^{2j-1}) - \mathfrak{L}'(h_N^{2j})\right| \\ &\le \|\mathfrak{L}'\|_{\infty, K} + C \|\mathfrak{L}''\|_{\infty, K} \le C_3, \end{aligned} \tag{4.84}$$

with $\|f\|_{\infty, K} = \sup_{x \in K} |f(x)|$ being the sup norm on the compact K. □

5. Decay rate of large area probability.

5.1. *Proof of Proposition* 2.3 (*Decay rate of large area probability*). We will display here the proof of Proposition 2.3 subject to Lemma 5.1, Corollary 5.2 and Lemmas 5.3, 5.4 that are stated and proven below.

In what follows, we use the notation $\|\mathbf{x}\| = \max\{|x_1|, |x_2|\}$ and $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2$ and $\mathrm{d}(\mathbf{x}, F) = \inf_{\mathbf{y} \in F} \|\mathbf{x} - \mathbf{y}\|$ for $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$ and $F \subset \mathbb{R}^2$. We also denote by ∂F the boundary of $F \subset \mathbb{R}^2$.

LEMMA 5.1. *For all $(j_1, j_2) \in (\mathbb{N} \cup \{0\})^2$ and all compact and convex subsets K in $\mathcal{D}$, there exists $c > 0$ such that*

$$\sup_{\mathbf{h} \in K} \left| \partial^{(j_1, j_2)} \left[\frac{1}{n} \mathfrak{L}_{\Lambda_n} \right](\mathbf{h}) - \partial^{(j_1, j_2)} \mathfrak{L}_\Lambda(\mathbf{h}) \right| \leq \frac{c}{n}, \qquad n \in \mathbb{N}. \tag{5.1}$$

PROOF. For all $(j_1, j_2) \in \mathbb{N}^2$, we first differentiate inside the integral

$$\partial^{(j_1, j_2)} \mathfrak{L}_\Lambda(\mathbf{h}) = \int_0^1 \partial_{h_0, h_1}^{(j_1, j_2)} \mathfrak{L}(x h_0 + h_1)\, dx. \tag{5.2}$$

Then, by using the error estimate for the Riemann sum of $x \mapsto \partial_{h_0, h_1}^{(j_1, j_2)} \mathfrak{L}(x h_0 + h_1)$, we obtain the result. □

By applying Lemma 5.1 for $(j_1, j_2) = (0, 1)$ and $(j_1, j_2) = (1, 0)$, we immediately obtain the following.

COROLLARY 5.2. *For all compact and convex subsets K in $\mathcal{D}$, there exist a $c > 0$ such that*

$$\sup_{\mathbf{h} \in K} \left\| \nabla \left[\frac{1}{n} \mathfrak{L}_{\Lambda_n} \right](\mathbf{h}) - \nabla \mathfrak{L}_\Lambda(\mathbf{h}) \right\| \leq \frac{c}{n}, \qquad n \in \mathbb{N}. \tag{5.3}$$

For $\eta > 0$, we let K_η be the compact and convex subset of $\mathcal{D}$ defined as

$$\begin{aligned} K_\eta := \Big\{ \mathbf{h} = (h_0, h_1) \in \mathbb{R}^2 \colon h_1 \in \Big[-\frac{\beta}{2} + \eta, \frac{\beta}{2} - \eta \Big], \\ h_0 + h_1 \in \Big[-\frac{\beta}{2} + \eta, \frac{\beta}{2} - \eta \Big] \Big\}. \end{aligned} \tag{5.4}$$

LEMMA 5.3. *The function $\nabla \mathfrak{L}_\Lambda \colon \mathcal{D} \mapsto \mathbb{R}^2$ defined as*

$$\begin{aligned} \nabla \mathfrak{L}_\Lambda(\mathbf{h}) &= (\partial_{h_0} \mathfrak{L}_\Lambda, \partial_{h_1} \mathfrak{L}_\Lambda)(\mathbf{h}) \\ &= \left(\int_0^1 x \mathfrak{L}'(x h_0 + h_1)\, dx, \int_0^1 \mathfrak{L}'(x h_0 + h_1)\, dx \right) \end{aligned} \tag{5.5}$$

is a $\mathcal{C}^1$ diffeomorphism. Moreover, for all $M > 0$ there exists a $\eta > 0$ such that $\|\nabla \mathfrak{L}_\Lambda(\mathbf{h})\| > M$ for $\mathbf{h} \in \mathcal{D} \setminus K_\eta$.

PROOF. The fact that $h \mapsto \mathfrak{L}'(h)$ is $\mathcal{C}^1$ and that $\mathfrak{L}''(h)$ is strictly positive on $(-\frac{\beta}{2}, \frac{\beta}{2})$ ensures that $\nabla \mathfrak{L}_\Lambda$ is $\mathcal{C}^1$ and that its Jacobian determinant that takes value

$$\begin{aligned} J_{\mathbf{h}} \nabla \mathfrak{L}_\Lambda = & \int_0^1 x^2 \mathfrak{L}''(xh_0 + h_1)\, dx \int_0^1 \mathfrak{L}''(xh_0 + h_1)\, dx \\ & - \left[\int_0^1 x \mathfrak{L}''(xh_0 + h_1)\, dx \right]^2 \end{aligned} \tag{5.6}$$

is, by Cauchy–Schwarz inequality, strictly positive. Thus, the proof that $\nabla \mathfrak{L}_\Lambda$ is a C^1 diffeomorphism from $\mathcal{D}$ to $\mathbb{R}^2$ will be complete once we show that $\nabla \mathfrak{L}_\Lambda$ is a bijection from $\mathcal{D}$ to $\mathbb{R}^2$.

At this stage, we note that for each $\mathbf{y} \in \mathbb{R}^2$ the function

$$T_{\mathbf{y}} : \mathbf{h} \to \mathfrak{L}_\Lambda(\mathbf{h}) - \mathbf{y} \cdot \mathbf{h} \tag{5.7}$$

is strictly convex and tends to ∞ as $\mathrm{d}(\mathbf{h}, \partial \mathcal{D}) \to 0$. Therefore, $T_{\mathbf{y}}$ admits a unique minimum on $\mathcal{D}$ at $\widetilde{\mathbf{h}}(\mathbf{y})$ that is also the unique solution of $\nabla \mathfrak{L}_\Lambda(\mathbf{h}) = \mathbf{y}$. Thus, $\nabla \mathfrak{L}_\Lambda$ is a bijection from $\mathcal{D}$ to $\mathbb{R}^2$.

We complete the proof of this lemma by assuming that there exists an $M_0 > 0$ and a sequence $(\mathbf{h}_n)_{n=0}^\infty$ in $\mathcal{D}$ so that $\mathrm{d}(\mathbf{h}_n, \partial \mathcal{D}) \to 0$ as $n \to \infty$ and $\|\nabla \mathfrak{L}_\Lambda(\mathbf{h}_n)\| \leq M_0$. Then set $\mathbf{y}_n = \nabla \mathfrak{L}_\Lambda(\mathbf{h}_n)$ and recall that $\mathbf{h}_n$ is the minimum of $T_{\mathbf{y}_n}$ for all $n \in \mathbb{N}$. However, $T_{\mathbf{y}_n}(0, 0) = 0$ and consequently $T_{\mathbf{y}_n}(\mathbf{h}_n) \leq 0$ for all $n \in \mathbb{N}$ and then $\mathfrak{L}_\Lambda(\mathbf{h}_n) \leq \mathbf{y}_n \cdot \mathbf{h}_n$ which brings a contradiction because $\lim_{n\to\infty} \mathfrak{L}_\Lambda(\mathbf{h}_n) = \infty$ [since $\mathrm{d}(\mathbf{h}_n, \partial \mathcal{D}) \to 0$] whereas $\mathbf{y}_n \cdot \mathbf{h}_n$ is smaller than M_0 times the diameter of $\mathcal{D}$. □

LEMMA 5.4. *For $n \in \mathbb{N} \setminus \{0, 1\}$, the function $\nabla[\frac{1}{n} \mathfrak{L}_{\Lambda_n}] : \mathcal{D}_n \mapsto \mathbb{R}^2$ defined as*

$$\nabla \left[\frac{1}{n} \mathfrak{L}_{\Lambda_n} \right](\mathbf{h}) = \partial_{h_0} \left[\frac{1}{n} \mathfrak{L}_{\Lambda_n}, \partial_{h_1} \mathfrak{L}_{\Lambda_n} \right](\mathbf{h}) \tag{5.8}$$

$$= \left(\frac{1}{n} \sum_{i=0}^{n-1} \frac{i}{n} \mathfrak{L}' \left(\frac{i}{n} h_0 + h_1 \right), \frac{1}{n} \sum_{i=0}^{n-1} \mathfrak{L}' \left(\frac{i}{n} h_0 + h_1 \right) \right) \tag{5.9}$$

is a $\mathcal{C}^1$ diffeomorphism. Moreover, for all $M > 0$ there exists a $\eta > 0$ and a $n_0 \in \mathbb{N}$ so that $\|\nabla[\frac{1}{n} \mathfrak{L}_{\Lambda_n}](\mathbf{h})\| > M$ for $n \geq n_0$ and $\mathbf{h} \in \mathcal{D}_n \setminus K_\eta$.

PROOF. The first part of the proof, that is, showing that $\nabla[\frac{1}{n} \mathfrak{L}_{\Lambda_n}]$ is a $\mathcal{C}^1$ diffeomorphism, is similar to that of Lemma 5.3 above. For the second part of the lemma, we first note that $\lim_{\eta \to 0^+} \min\{\mathfrak{L}_\Lambda(\mathbf{h}) : \mathbf{h} \in \partial K_\eta\} = \infty$. Then, for a given $M > 0$, we can pick $\eta_0 > 0$ so that $\mathfrak{L}_\Lambda$ remains larger than $2M$ on ∂K_{η_0}. Moreover, Lemma 5.1 ensures that $\frac{1}{n} \mathfrak{L}_{\Lambda_n}$ converges to $\mathfrak{L}_\Lambda$ uniformly on K_{η_0} and, therefore, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\frac{1}{n} \mathfrak{L}_{\Lambda_n}$ remains strictly larger than M on ∂K_{η_0}. Consider $\mathbf{h} \in \mathcal{D}_n \setminus K_\eta$ and let $t \in (0, 1)$ be the unique solution of $t\mathbf{h} \in \partial K_{\eta_0}$.

By convexity and since $\frac{1}{n}\mathfrak{L}_{\Lambda_n}(0,0)=0$, we claim that $\frac{1}{n}\mathfrak{L}_{\Lambda_n}(\mathbf{h}) \geq \frac{1}{n}\mathfrak{L}_{\Lambda_n}(t\mathbf{h}) > M$ which completes the proof. □

REMARK 5.5. As in the proof of Lemma 5.3 above, we denote by $\widetilde{\mathbf{h}} := (\widetilde{h}_0, \widetilde{h}_1)$ the inverse function of $\nabla\mathfrak{L}_\Lambda$. Since $\mathfrak{L}$ is an even function, we easily obtain, for instance, by observing that $T_{(q,0)}(h_0,h_1) = T_{(q,0)}(h_0, -h_0-h_1)$, that $\widetilde{h}_0(q,0) = -2\widetilde{h}_1(q,0) > 0$ for all $q>0$. We will also denote by $\mathbf{h}_n^q := (h_{n,0}^q, h_{n,1}^q)$ the unique solution of $\nabla[\frac{1}{n}\mathfrak{L}_{\Lambda_n}](\mathbf{h}) = (q,0)$ for all $n \geq 2$ and $q>0$. Again the fact that $\mathfrak{L}$ is even ensures that $h_{n,0}^q(1-\frac{1}{n}) = -2h_{n,1}^q > 0$.

At this stage, we have enough tools to prove Proposition 2.3.

PROOF OF PROPOSITION 2.3. Pick $q \in [q_1,q_2]$, $n \in \mathbb{N}$ and note that

$$\left|\left[\frac{1}{n}\mathfrak{L}_{\Lambda_n}(\mathbf{h}_n^q) - h_{n,0}^q q\right] - \left[\mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q,0)) - \widetilde{h}_0(q,0)q\right]\right| \leq A+B+C \tag{5.10}$$

with

$$\begin{aligned} A &= \left|\frac{1}{n}\mathfrak{L}_{\Lambda_n}(\mathbf{h}_n^q) - \mathfrak{L}_\Lambda(\mathbf{h}_n^q)\right|, \\ B &= \left|\mathfrak{L}_\Lambda(\mathbf{h}_n^q) - \mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q,0))\right|, \qquad C = q\left|h_{n,0}^q - \widetilde{h}_0(q,0)\right|. \end{aligned} \tag{5.11}$$

From Lemma 5.4, we know that there exists an $\eta>0$ and a $n_0 \in \mathbb{N}$ such that $\mathbf{h}_n^q \in K_\eta$ for all $q \in [q_1,q_2]$ and $n \geq n_0$. By using Lemma 5.1 with $(j_1,j_2)=(0,0)$ and $K=K_\eta$, we can claim that there exists a $c_1>0$ satisfying $A \leq \frac{c_1}{n}$ for $n \geq n_0$ and $q \in [q_1,q_2]$. The B quantity is dealt with by applying Corollary 5.2 with $K=K_\eta$, that is there exists a $c_2>0$ such that

$$\sup_{x \in K_\eta}\left\|\nabla\left[\frac{1}{n}\mathfrak{L}_{\Lambda_n}\right](x) - \nabla\mathfrak{L}_\Lambda(x)\right\| \leq \frac{c_2}{n}, \qquad n \geq n_0. \tag{5.12}$$

Therefore, for $q \in [q_1,q_2]$ and $n \geq n_0$ we can write

$$\begin{aligned} \nabla\left[\frac{1}{n}\mathfrak{L}_{\Lambda_n}\right](\mathbf{h}_n^q) &= \nabla\mathfrak{L}_\Lambda(\mathbf{h}_n^q) + \varepsilon_{n,q}, \\ (q,0) &= \nabla\mathfrak{L}_\Lambda(\mathbf{h}_n^q) + \varepsilon_{n,q} \end{aligned} \tag{5.13}$$

with $\|\varepsilon_{n,q}\| \leq \frac{c_2}{n}$. Therefore, by Lemma 5.3, we can claim that $\mathbf{h}_n^q = \widetilde{\mathbf{h}}((q,0)-\varepsilon_{n,q})$. We set

$$Q_n = \left\{(x,y) \in \mathbb{R}^2 : \mathrm{d}\big((x,y),[q_1,q_2]\times\{0\}\big) \leq \frac{c_2}{n}\right\},$$

so that there exists a $n_1 \geq n_0$ such that Q_{n_1} is a convex subset of $\mathcal{D}$ and since $\mathbf{x} \mapsto \widetilde{\mathbf{h}}(\mathbf{x})$ is $\mathcal{C}^1$ on $\mathcal{D}$ we can claim that $\widetilde{\mathbf{h}}$ is Lipschitz on Q_{n_1}. Thus, there exists a $c_3>0$ such that

$$\left\|\mathbf{h}_n^q - \widetilde{\mathbf{h}}\big((q,0)\big)\right\| \leq c_3\|\varepsilon_{n,q}\| \leq \frac{c_2c_3}{n}, \qquad q \in [q_1,q_2], n \geq n_1, \tag{5.14}$$

and this proves (2.29). Moreover,

$$C \le q_2 \|\mathbf{h}_n^q - \widetilde{\mathbf{h}}((q,0))\| \le \frac{q_2 c_2 c_3}{n}, \qquad q \in [q_1, q_2], n \ge n_1. \tag{5.15}$$

Finally, since $\mathfrak{L}_\Lambda$ is $\mathcal{C}^1$ on $\mathcal{D}$, there exists a $c_4 > 0$ such that $\mathfrak{L}_\Lambda$ is Lipschitz with constant c_4 on Q_{n_1}. Thus,

$$B \le c_4 \|\mathbf{h}_n^q - \widetilde{\mathbf{h}}((q,0))\| \le \frac{c_2 c_3 c_4}{n}, \qquad q \in [q_1, q_2], n \ge n_1. \tag{5.16}$$

This completes the proof of Proposition 2.3. □

6. Limit theorems for the joint distribution. In Section 6.1 below, we give a proof of Proposition 2.2 which estimates, uniformly in $q \in [q_1, q_2] \subset (0, \infty)$, the probability of the event $\{\Lambda_n = (Y_n, V_n) = (nq, 0)\}$ under the tilted law $\mathbf{P}_{n,\mathbf{h}_n^q}$ [recall (2.26)]. To that aim, we state and prove Proposition 6.1, which gives a local central limit theorem for (Y_n, V_n) under $\mathbf{P}_{n,\mathbf{h}_n^q}$. In Section 6.2, we prove Proposition 2.5 which allows us to bound from below the probability that, under $\mathbf{P}_\beta$ and conditioned on both $V_n = 0$ and $Y_n = nq$ the random walk V remains strictly positive.

6.1. *Proof of Proposition* 2.2. We display the proof of Proposition 2.2 which turns out to be a straightforward consequence of Proposition 6.1 below. The latter proposition will be proven at the end of the section.

PROOF. Recall (2.21)–(2.26) and for any $\mathbf{h} \in \mathcal{D}$, define the matrix

$$\mathbf{B}(\mathbf{h}) := \operatorname{Hess} \mathfrak{L}_\Lambda(\mathbf{h}) \tag{6.1}$$

and let Θ be the Gaussian random vector with zero mean and covariance matrix $\mathbf{B}(\mathbf{h})$. We denote the density of Θ by

$$f_{\mathbf{h}}(X) = \frac{1}{2\pi\sqrt{\det \mathbf{B}(\mathbf{h})}} \exp\left(-\frac{1}{2}\langle \mathbf{B}(\mathbf{h})^{-1} X, X\rangle\right), \qquad X \in \mathbb{R}^2, \tag{6.2}$$

and its characteristic function by

$$\bar{\Phi}_{\mathbf{h}}(T) = \exp\left(-\tfrac{1}{2}\langle \mathbf{B}(\mathbf{h}) T, T\rangle\right), \qquad T \in \mathbb{R}^2. \tag{6.3}$$

Consider now the case $(Y_N, V_N) = (N q_{N,L}, 0)$ as in Section 4.2 and recall that $q_{N,L} \in [\frac{1}{2a_2^2}, \frac{1}{a_1^2}]$. We will show that the local central limit theorem below is valid uniformly in q in some compact subsets.

PROPOSITION 6.1. *For* $[q_1, q_2] \subset \mathbb{R}$, *we have* $\lim_{N \to +\infty} \tau_N = 0$ *with*

$$\begin{aligned} \tau_N := \sup_{q \in [q_1, q_2]} \sup_{x, y \in \mathbb{Z}} \Big| & N^2 \mathbf{P}_{N, \mathbf{h}_N^q}(N Y_N = N^2 q + x, V_N = y) \\ & - f_{\widetilde{\mathbf{h}}(q,0)}\left(\frac{x}{N^{3/2}}, \frac{y}{\sqrt{N}}\right)\Big|. \end{aligned} \tag{6.4}$$

By applying Proposition 6.1 with $x = y = 0$, we obtain that

$$\sup_{q\in[q_1,q_2]} \left|N^2 \mathbf{P}_{N,\mathbf{h}_N^q}(NY_N = N^2 q, V_N = 0) - f_{\widetilde{\mathbf{h}}(q,0)}(0,0)\right| \le \tau_N \to 0, \tag{6.5}$$

and since the Hessian matrix $\mathbf{B}(\widetilde{\mathbf{h}}(q,0))$ is uniformly bounded in $q \in [q_1, q_2]$, we observe that there exists $C > 0$ such that

$$\frac{1}{CN^2} \le \mathbf{P}_{N,\mathbf{h}_N^q}(NY_N = N^2 q, V_N = 0) \le \frac{C}{N^2} \qquad \text{for } N \text{ large enough}, \tag{6.6}$$

which completes the proof of Proposition 2.2. □

6.1.1. *Proof of Proposition* 6.1. We follow closely the proof of Dobrushin and Hryniv (1996), making sure that the result holds uniformly in $q \in [q_1, q_2]$. From Lemmas 5.3 and 5.4, there exists $\eta > 0$ such that both $\widetilde{\mathbf{h}}(q,0)$ and $\mathbf{h}_N^q$ are in K_η for all $q \in [q_1, q_2]$ and for N large enough.

We let $\mathfrak{E}(z) = \mathbf{E}_\beta(e^{zU_1})$ be the holomorphic function defined on the strip $\{z \in \mathbb{C}\colon \mathrm{Re}(z) \in (-\beta/2, \beta/2)\}$. For any $h \in (-\beta/2, \beta/2)$ and $t \in \mathbb{R}$, we set

$$\varphi_h(t) := \mathfrak{E}(h + it)/\mathfrak{E}(h). \tag{6.7}$$

Let us state some properties of the function $\varphi_h(t)$ that will be used in the sequel [they are established in Dobrushin and Hryniv (1996)]. First of all, for any $h \in \mathcal{K} := [-\beta/2 + \eta, \beta/2 - \eta]$ and $t \in \mathbb{R}$

$$\left|\varphi_h(t)\right| \le \varphi_h(0) = 1. \tag{6.8}$$

Second, for any $\delta \in (0, \pi)$, there exists a constant $C = C(\mathcal{K}, \delta) > 0$ such that for every $h \in \mathcal{K}$ and any $t \in [\delta, 2\pi - \delta]$, we have

$$\left|\varphi_h(t)\right| \le e^{-C}. \tag{6.9}$$

And finally, there exists a constant $\alpha = \alpha(\mathcal{K}) > 0$ such that for all $h \in \mathcal{K}$ and any t, $|t| \le \pi$, the following inequality holds:

$$\left|\varphi_h(t)\right| \le \exp\left(-\alpha^2 t^2 \mathfrak{L}''(h)\right). \tag{6.10}$$

For any $T = (t_0, t_1) \in \mathbb{R}^2$, let $\Phi_{N,\mathbf{h}_N^q}(T)$ be the characteristic function of the random vector $\Lambda_N = (Y_N, V_N)$. Let us rewrite it with the functions $\varphi_h(t)$,

$$\Phi_{N,\mathbf{h}_N^q}(T) = \mathbf{E}_{N,\mathbf{h}_N^q}\left[e^{i\langle T,\Lambda_N\rangle}\right] = \prod_{j=1}^{N} \varphi_{h_{j,N}}(t_{j,N}), \tag{6.11}$$

where

$$h_{j,N} = \left(1 - \frac{j}{N}\right) h_{N,0}^q + h_{N,1}^q \quad \text{and} \quad t_{j,N} = \left(1 - \frac{j}{N}\right) t_0 + t_1. \tag{6.12}$$

Note that

$$\hat{\Phi}_{N,\mathbf{h}_N^q}(T) = \Phi_{N,\mathbf{h}_n^q}(N^{-1/2}T) \exp\left(-\frac{i}{\sqrt{N}} \left\langle T, \mathbf{E}_{N,\mathbf{h}_N^q}(\Lambda_N)\right\rangle\right) \tag{6.13}$$

is the characteristic function of the centered random vector $\Lambda_N^\star := \Lambda_N - \mathbf{E}_{N,\mathbf{h}_N^q}(\Lambda_N)$.

Let $\mathbf{v}_N = (\frac{x}{N^{3/2}}, \frac{y}{\sqrt{N}})$. Using the well know inversion formula for the Fourier transform, we rewrite the LHS of (6.4), that is,

$$R_N = N^2 \mathbf{P}_{N,\mathbf{h}_N^q}\big(NY_N = N^2 q + x, V_N = y\big) - f_{\widetilde{\mathbf{h}}(q,0)}(\mathbf{v}_N) \tag{6.14}$$

in the form

$$\begin{aligned} R_N = {} & \frac{1}{(2\pi)^2} \int_{\mathcal{A}} \hat{\Phi}_{N,\mathbf{h}_N^q}(T) e^{-i\langle T, \mathbf{v}_N \rangle} \, dT \\ & - \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \bar{\Phi}_{\widetilde{\mathbf{h}}(q,0)}(T) e^{-i\langle T, \mathbf{v}_N \rangle} \, dT, \end{aligned} \tag{6.15}$$

where

$$\mathcal{A} = \big\{ T = (t_0, t_1) \in \mathbb{R}^2 \colon |t_0| \le \pi N^{3/2}, |t_1| \le \pi \sqrt{N} \big\}. \tag{6.16}$$

Following the proof in Dobrushin and Hryniv (1996), we bound the LHS of (6.15) by the sum of four terms,

$$|R_N| \le (2\pi)^{-2} \big(J_1^{(q)} + J_2^{(q)} + J_3^{(q)} + J_4^{(q)} \big), \tag{6.17}$$

where, for some positive constants A and Δ,

$$J_1^{(q)} = \int_{\mathcal{A}_1} \big| \hat{\Phi}_{N,\mathbf{h}_N^q}(T) - \bar{\Phi}_{\widetilde{\mathbf{h}}(q,0)}(T) \big| \, dT, \qquad \mathcal{A}_1 = [-A, A]^2, \tag{6.18}$$

$$J_2^{(q)} = \int_{\mathcal{A}_2} \bar{\Phi}_{\widetilde{\mathbf{h}}(q,0)}(T) \, dT, \qquad \mathcal{A}_2 = \mathbb{R}^2 \setminus \mathcal{A}_1, \tag{6.19}$$

$$\begin{aligned} J_3^{(q)} = {} & \int_{\mathcal{A}_3} \big| \hat{\Phi}_{N,\mathbf{h}_N^q}(T) \big| \, dT, \\ & \qquad \mathcal{A}_3 = \big\{ T \in \mathbb{R}^2 \colon |t_l| \le \Delta \sqrt{N}, l = 0, 1 \big\} \setminus \mathcal{A}_1, \end{aligned} \tag{6.20}$$

$$J_4^{(q)} = \int_{\mathcal{A}_4} \big| \hat{\Phi}_{N,\mathbf{h}_N^q}(T) \big| \, dT, \qquad \mathcal{A}_4 = \mathcal{A} \setminus (\mathcal{A}_1 \cup \mathcal{A}_3). \tag{6.21}$$

For an arbitrary $\varepsilon > 0$, Dobrushin and Hryniv proved that for a convenient choice of the constants $A = A(\varepsilon)$ and Δ, we have the bounds $J_i^{(q)} < \varepsilon/4$ for $i = 1, 2, 3, 4$ for sufficiently large N. Therefore, the proof will be complete once we show that this assertion is also valid uniformly in $q \in [q_1, q_2]$. It remains to evaluate all $J_i^{(q)}$.

First, we bound $J_1^{(q)}$. For $\mathbf{h} \in \mathcal{D}_n$, define the matrix

$$\mathbf{B}_n(\mathbf{h}) := \frac{1}{n} \operatorname{Hess} \mathfrak{L}_{\Lambda_n}(\mathbf{h}), \qquad n \in \mathbb{N}. \tag{6.22}$$

By Lemma 5.1 and Proposition 2.3, we obtain the relation

$$\mathbf{B}_N(\mathbf{h}_N^q) = \mathbf{B}(\widetilde{\mathbf{h}}(q,0)) + R'_N, \tag{6.23}$$

with the bound $|R'_N| \le C_1(q_1, q_2)N^{-1}$ uniform in $q \in [q_1, q_2]$.

Recall that $\mathcal{K} = [-\beta/2 + \eta, \beta/2 - \eta]$. Since $\mathfrak{E}$ is holomorphic on $\{z \in \mathbb{C}: \mathrm{Re}(z) \in (-\beta/2, \beta/2)\}$, for any $\eta > 0$ there exists an $A' > 0$ so that $\mathrm{Re}(\mathfrak{E}(z)) > 0$ for $z \in \mathcal{K} + i[-A', A']$ and, therefore, we can use a branch of the complex logarithm to extend the function $\mathfrak{L}$ (that equals $\log \mathfrak{E}$) to $\mathcal{K} + i[-A', A']$. We observe that $\mathbf{h} \in K_\eta$ and $T \in \frac{1}{2}[-A', A']^2$ yield $(1 - \frac{j}{n})h_0 + h_1 \in \mathcal{K}$ and $(1 - \frac{j}{n})t_0 + t_1 \in [-A', A']$ for all $j \in \{1, \dots, N\}$. Thus, we can extend $\mathfrak{L}_{\Lambda_n}$ to $K_\eta \times \frac{1}{2}[-A', A']^2$ with the formula

$$\mathfrak{L}_{\Lambda_n}(\mathbf{h} + iT) := \sum_{j=1}^{n} \mathfrak{L}\Big(\Big(1 - \frac{j}{n}\Big)(h_0 + it_0) + h_1 + it_1\Big). \tag{6.24}$$

Similarly, we extend $\mathfrak{L}_\Lambda$ to $K_\eta \times \frac{1}{2}[-A', A']^2$ and Lemma 5.1 can, without further difficulty, be extended to $K_\eta \times \frac{1}{2}[-A', A']^2$. In particular, any partial derivative of order 3 of $\frac{1}{n}\mathfrak{L}_{\Lambda_n}$ converges uniformly to its counterpart of $\mathfrak{L}_\Lambda$ on $K_\eta \times \frac{1}{2}[-A', A']^2$. Consequently, for N large enough, we make sure that for $q \in [q_1, q_2]$ and for $T \in \mathcal{A}_1$, we have $\mathbf{h}_N^q \in K_\eta$ and $T/N \in \frac{1}{2}[-A', A']^2$ so that we can consider the remainder

$$\begin{aligned} R''_N = {} & \mathfrak{L}_{\Lambda_N}(\mathbf{h}_N^q + iN^{-1/2}T) - \mathfrak{L}_{\Lambda_N}(\mathbf{h}_N^q) - \frac{i}{\sqrt{N}}\langle T, \mathbf{E}_{N,\mathbf{h}_N^q}(\Lambda_N)\rangle \\ & + \frac{1}{2}\langle \mathbf{B}_N(\mathbf{h}_N^q)T, T\rangle, \end{aligned} \tag{6.25}$$

and apply a Taylor–Lagrange inequality to assert that there exists a constant $C(A, q_1, q_2) > 0$ such that for N large enough $|R'_N| \le C(A, q_1, q_2)/\sqrt{N}$ uniformly in $q \in [q_1, q_2]$ and $T \in \mathcal{A}_1$.

Therefore, we can use (6.3), (6.7), (6.11)–(6.13) and (6.23) to get, as $N \to +\infty$,

$$\begin{aligned} & \sup_{q \in [q_1, q_2], T \in \mathcal{A}_1} \big|\hat{\Phi}_{N,\mathbf{h}_N^q}(T) - \bar{\Phi}_{\widetilde{\mathbf{h}}(q,0)}(T)\big| \\ & = \sup_{q \in [q_1, q_2], T \in \mathcal{A}_1} \big|e^{(1/2)R'_N\|T\|^2 + R''_N} - 1\big| \to 0. \end{aligned} \tag{6.26}$$

Hence, for every finite $A > 0$, we obtain the convergence $J_1^{(q)} \to 0$ as $N \to \infty$ uniformly in $q \in [q_1, q_2]$.

Let $\underline{B}$ be such that $0 < \underline{B} \le \mathbf{B}(\widetilde{\mathbf{h}}(q, 0))$ for all $q \in [q_1, q_2]$. Hence, we can bound $J_2^{(q)}$ as follows:

$$\sup_{q \in [q_1, q_2]} J_2^{(q)} \le \int_{\mathcal{A}_2} e^{-(1/2)\langle \underline{B}T, T\rangle}\, dT \to 0 \qquad \text{as } A \to \infty. \tag{6.27}$$

To estimate $J_3^{(q)}$, we fix any $T \in \mathcal{A}_3$ and put $\Delta = \pi/2$. Then all the numbers $t_{j,N}$ in (6.12) satisfy the condition $|t_{j,N}| \le \pi\sqrt{N}$, evaluating each factor in (6.11) with the help of (6.10) and (6.23) we obtain the bound

$$(6.28)\quad \left|\hat{\Phi}_{N,\mathbf{h}_N^q}(T)\right| \le \exp\bigl(-\alpha^2\langle \mathbf{B}_N(\mathbf{h}_N^q)T, T\rangle\bigr) \le C\exp\bigl(-\alpha^2\langle \mathbf{B}(\widetilde{\mathbf{h}}(q,0))T, T\rangle\bigr),$$

for some constant $C > 0$. As a result, as $A \to \infty$,

$$(6.29)\quad \begin{aligned} \sup_{q\in[q_1,q_2]} J_3^{(q)} &= \sup_{q\in[q_1,q_2]} \int_{\mathcal{A}_3} \left|\hat{\Phi}_{N,\mathbf{h}_N^q}(T)\right| dT \\ &\le C\int_{\mathcal{A}_2} \exp\bigl(-\alpha\langle \underline{B}T, T\rangle\bigr)\, dT \to 0. \end{aligned}$$

To evaluate $J_4^{(q)}$ put $\delta = \frac{1}{17(2)^2}$ and for any $T \in \mathcal{A}_4$ denote by $\mathbf{N}_N(T)$ the number of indexes $j = 1, 2, \dots, N$ such that $\tau_{j,N} \notin \mathcal{O}_\delta := \bigcup_{m\in\mathbb{Z}}[m-\delta, m+\delta]$, where

$$(6.30)\quad \tau_{j,N} := \frac{1}{2\pi\sqrt{N}} t_{j,N}.$$

Use (6.8) and (6.9) to estimate those factors in (6.11) and we have

$$(6.31)\quad \left|\hat{\Phi}_{N,\mathbf{h}_N^q}(T)\right| = \prod_{j=1}^{N} \left|\varphi_{h_{j,N}}\left(\frac{1}{\sqrt{N}} t_{j,N}\right)\right| \le \exp\bigl(-C\mathbf{N}_N(T)\bigr).$$

A lower bound of $\mathbf{N}_N(T)$ is given in Dobrushin and Hryniv (1996), page 443: for all $T \in \mathcal{A}_4$ and N large enough, there exists a constant $\kappa > 0$ such that $\mathbf{N}_N(T) \ge \kappa N$. Then, uniformly in $q \in [q_1, q_2]$,

$$(6.32)\quad J_4^{(q)} = \int_{\mathcal{A}_4} \left|\hat{\Phi}_{N,\mathbf{h}_N^q}(T)\right| dT \le (2\pi)^2 N^2 \exp(-C\kappa N) \to 0 \qquad \text{as } N \to \infty.$$

6.2. *Proof of Proposition* 2.5 (*Unique excursion for large area*). From now on, the letters $C, C', C_1, \dots$ shall denote constants that do not depend on N and on $q \in [q_1, q_2] \subset (0, \infty)$. In other words, all the bounds we are going to establish are uniform in $N \ge N_0$ and $q \in [q_1, q_2]$.

To begin with, we prove Lemma 6.4 subject to Lemmas 6.2 and 6.3 below. Lemma 6.4 is crucial in the proof of Proposition 2.5. It allows us indeed to bound from below, for any $j \in \mathbb{N}$, the probability that the random walk V, conditioned on making a large area, is below 0 at time j. Such a lower bound was available in Dobrushin and Hryniv (1996) but only for j of order N. Here, we deal with any $j \le N$. The first step of the proof is an upper bound on the moment generating function of the tilted random walk V.

LEMMA 6.2. *There exist three positive constants C', C_1, λ such that for every integer $j \le N/2$, the following bound holds*:

$$(6.33)\quad \mathbf{E}_{N,\mathbf{h}_N^q}\bigl[e^{-\lambda V_j}\bigr] \le C' e^{-C_1 j}, \qquad N \in \mathbb{N}.$$

PROOF. Under the tilted law [see (2.24)] the increments $U_i = V_i - V_{i-1}$ are still independent but no more identically distributed. For any positive λ, we have

$$\log \mathbf{E}_{N,\mathbf{h}_N^q}[e^{-\lambda V_j}] = \sum_{1\le i\le j} \big(\mathfrak{L}(-\lambda + h_N^i) - \mathfrak{L}(h_N^i)\big) \tag{6.34}$$

with $h_N^i := (1-\frac{i}{N})h_{N,0}^q + h_{N,1}^q$. By Remark 5.5, we know that for all $q>0$ and $N \ge 2$,

$$h_{N,0}^q\Big(1-\frac{1}{N}\Big) = -2h_{N,1}^q > 0. \tag{6.35}$$

A straightforward consequence of (6.35) is that $h_N^i \ge 0$ for all $i \le N/2$. Then the convexity of $\mathfrak{L}(\cdot)$ and the fact that $\mathfrak{L}(0) = \mathfrak{L}'(0) = 0$ yield that there exists a $c>0$ so that for all $i \le N/2$ and λ small enough

$$\mathfrak{L}(-\lambda + h_N^i) - \mathfrak{L}(h_N^i) \le \mathfrak{L}(-\lambda) \le c\lambda^2. \tag{6.36}$$

We established in Proposition 2.3 the existence of $C>0$ and $N_0 \in \mathbb{N}$ such that for all $N \ge N_0$, and every $q \in [q_1, q_2]$, we have

$$\|\mathbf{h}_N^q - \widetilde{\mathbf{h}}(q,0)\| \le \frac{C}{N}. \tag{6.37}$$

Thanks to Lemma 5.3 and Remark 5.5, there exists a constant $R>0$ such that

$$\widetilde{h}_0(q,0) \ge R > 0 \qquad \forall q \in [q_1, q_2]. \tag{6.38}$$

Thus, provided N_0 is chosen large enough, we deduce from (6.37) and (6.38) that $h_{N,0}^q \ge R/2$ for $N \ge N_0$ and $q \in [q_1,q_2]$. Moreover, thanks to (6.35), we also write $h_N^i \ge \frac{1}{4} h_{N,0}^q$ for $i \le N/4$ such that finally $h_N^i \ge R/8$ for $i \le N/4$. Observe that by convexity of $\mathfrak{L}(.)$,

$$\sum_{1\le i\le j} \big(\mathfrak{L}(-\lambda + h_N^i) - \mathfrak{L}(h_N^i)\big) \le -\lambda \sum_{1\le i\le j} \mathfrak{L}'(-\lambda + h_N^i). \tag{6.39}$$

Hence, for $j \le N/4$ and for $\lambda \le R/16$ we have

$$\sum_{1\le i\le j} \big(\mathfrak{L}(-\lambda + h_N^i) - \mathfrak{L}(h_N^i)\big) \le -\lambda j\, \mathfrak{L}'\Big(\frac{R}{16}\Big). \tag{6.40}$$

For $N/4 \le j \le N/2$ in turn we split the sum in the LHS of (6.40) into a sum over $i \le N/4$ [that is dealt with as in (6.40)] and a sum over $i \ge N/4$ [that is dealt with by using (6.36)]. Thus,

$$\begin{aligned}\sum_{1\le i\le j} \big(\mathfrak{L}(-\lambda + h_N^i) - \mathfrak{L}(h_N^i)\big) &= -\lambda \frac{N}{4}\mathfrak{L}'\Big(\frac{R}{16}\Big) + c\Big(j - \frac{N}{4}\Big)\lambda^2 \\ &\le \frac{N}{4}\Big(c\lambda^2 - \lambda \mathfrak{L}'\Big(\frac{R}{16}\Big)\Big).\end{aligned} \tag{6.41}$$

It remains to choose $\lambda > 0$ small enough to make sure that $c\lambda^2 - \lambda\mathfrak{L}'(R/16) > 0$ and then, (6.40) and (6.41) complete the proof. □

The next lemma ensures that we can restrict ourselves to $j \le N/2$.

LEMMA 6.3. *For $a \in \mathbb{R}$ and $j \in \{1, \dots, N\}$*

$$\begin{aligned}&\mathbf{P}_\beta(V_j \le a, Y_N = Nq, V_N = 0)\\&\qquad = \mathbf{P}_\beta(V_{N-j} \le a, Y_N = Nq, V_N = 0).\end{aligned} \tag{6.42}$$

PROOF. We just need to use time reversal, that is,

$$(V_N - V_{N-j}, 0 \le j \le N) \overset{d}{=} (V_j, 0 \le j \le N), \tag{6.43}$$

to obtain that

$$\begin{aligned}&\mathbf{P}_\beta(V_j \le a, Y_N = Nq, V_N = 0)\\&\qquad = \mathbf{P}_\beta(-V_{N-j} \le -a, -Y_N = Nq, V_N = 0).\end{aligned} \tag{6.44}$$

By using the symmetry of V, we complete the proof:

$$(-V_j, 0 \le j \le N) \overset{d}{=} (V_j, 0 \le j \le N). \tag{6.45}$$ □

At this stage, we need to use precise results for the local central limit theorem. We recall (2.26) and for convenience we use the notation

$$\begin{aligned}&\alpha_N^q := \mathbf{P}_{N,\mathbf{h}_N^q}(NY_N = N^2 q, V_N = 0) \quad \text{and}\\&\xi_N^q := \exp(\mathfrak{L}_{\Lambda_N}(\mathbf{h}_N^q) - N h_{N,0}^q q).\end{aligned} \tag{6.46}$$

Hence, we have

$$\mathbf{P}_\beta(Y_N = Nq, V_N = 0) = \xi_N^q \alpha_N^q. \tag{6.47}$$

We can handle α_N^q with the help of Proposition 2.2: there exists a $C_2 > 0$ such that

$$\frac{1}{C_2}\frac{1}{N^2} \le \alpha_N^q \le \frac{C_2}{N^2}. \tag{6.48}$$

Proposition 2.3 allows us to write that there exists a positive constant C_3 so that

$$e^{-C_3} e^{N(\mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q,0)) - \widetilde{h}_0(q,0)q)} \le \xi_N^q \le e^{C_3} e^{N(\mathfrak{L}_\Lambda(\widetilde{\mathbf{h}}(q,0)) - \widetilde{h}_0(q,0)q)}. \tag{6.49}$$

We can state the following.

LEMMA 6.4. *There exists a constant $\lambda > 0$ such that for all $a > 0, q \in [q_1, q_2]$, $N \ge N_0$ and $0 \le j \le N$*

$$\mathbf{P}_\beta(V_j \le -a, Y_N = Nq, V_N = 0) \le \xi_N^q C' e^{-C_1(j \wedge (N-j)) - \lambda a}. \tag{6.50}$$

PROOF. By the symmetry in Lemma 6.3, we can without loss of generality assume $j \le N/2$. By using Lemma 6.2, we can write

$$\begin{aligned}
&\mathbf{P}_\beta(V_j \le -a, Y_N = Nq, V_N = 0)\\
&\quad \le \mathbf{E}_\beta[e^{-\lambda V_j}, Y_N = Nq, V_N = 0]e^{-\lambda a}\\
&\quad = \xi_N^q e^{-\lambda a}\mathbf{E}_{N,\mathbf{h}_N^q}[e^{-\lambda V_j}, Y_N = Nq, V_N = 0]\\
&\quad \le \xi_N^q e^{-\lambda a}\mathbf{E}_{N,\mathbf{h}_N^q}[e^{-\lambda V_j}] \le \xi_N^q C' e^{-C_1 j - \lambda a}. \qquad \square
\end{aligned}$$

PROOF OF PROPOSITION 2.5. Let $u_N = \lfloor \nu \log N \rfloor$ where $\nu > 0$ will be chosen afterward. The first step is to write

$$\begin{aligned}
&\mathbf{P}_\beta(V_i > 0, 0 < i < N; NY_N = N^2 q, V_N = 0)\\
&\quad \ge \mathbf{P}_\beta(V_1 = V_{N-1} = u_N, V_i > 0, 2 < i < N-2;\\
&\qquad NY_N = N^2 q, V_N = 0).
\end{aligned} \tag{6.51}$$

By using Markov's property at time 1 and $N-1$, we obtain

$$\begin{aligned}
&\mathbf{P}_\beta(V_1 = V_{N-1} = u_N, V_i > 0, 2 < i < N-2; NY_N = N^2 q, V_N = 0)\\
&\quad = \mathbf{P}_\beta(U_1 = u_N)^2 \mathbf{P}_\beta(V_i > -u_N, 1 < i < N-3;\\
&\qquad (N-2)Y_{N-2} = N^2 q - (N-1)u_N, V_{N-2} = 0).
\end{aligned} \tag{6.52}$$

We shall use a basic lower bound

$$\begin{aligned}
&\mathbf{P}_\beta(V_i > -u_N, 1 < i < N-3; (N-2)Y_{N-2} = N^2 q - (N-1)u_N, V_{N-2} = 0)\\
&\ge \mathbf{P}_\beta((N-2)Y_{N-2} = N^2 q - (N-1)u_N, V_{N-2} = 0)\\
&\quad - \sum_{i=1}^{N-3} \mathbf{P}_\beta(V_i \le -u_N, (N-2)Y_{N-2} = N^2 q - (N-1)u_N, V_{N-2} = 0).
\end{aligned} \tag{6.53}$$

We take care of the second term by letting $q' = \frac{N^2 q - (N-1)u_N}{(N-2)^2}$ in Lemma 6.4

$$\begin{aligned}
&\sum_{i=1}^{N-3} \mathbf{P}_\beta(V_i \le -u_N, (N-2)Y_{N-2} = (N-2)^2 q', V_{N-2} = 0)\\
&\quad \le \xi_{N-2}^{q'} \sum_{i=1}^{N-3} C' e^{-C_1(i \wedge (N-2-i)) - \lambda u_N} \le C_4 \xi_{N-2}^{q'} e^{-\lambda u_N}.
\end{aligned} \tag{6.54}$$

Observe that thanks to the notation (6.46) we can write the first term in the RHS of (6.53) as

$$\mathbf{P}_\beta((N-2)Y_{N-2} = (N-2)^2 q', V_{N-2} = 0) = \xi_{N-2}^{q'} \alpha_{N-2}^{q'}. \tag{6.55}$$

Hence,

$$\begin{aligned} &\mathbf{P}_\beta\big(V_i > 0, 0 < i < N; NY_N = N^2 q, V_N = 0\big) \\ &\quad \geq \mathbf{P}_\beta(U_1 = u_N)^2 \xi_{N-2}^{q'}\big[\alpha_{N-2}^{q'} - C_4 e^{-\lambda u_N}\big]. \end{aligned} \tag{6.56}$$

Observe that

$$\mathbf{P}_\beta(U_1 = u_N)^2 = \frac{1}{c_\beta^2} e^{-\beta \lfloor \nu \log N \rfloor} \geq \frac{1}{c_\beta^2} N^{-\beta\nu} \tag{6.57}$$

and recall that $\mathbf{P}_\beta(NY_N = N^2 q, V_N = 0) = \xi_N^q \alpha_N^q$. Therefore,

$$\begin{aligned} &\mathbf{P}_\beta\big(V_i > 0, 0 < i < N | NY_N = N^2 q, V_N = 0\big) \\ &\quad \geq \frac{N^{-\beta\nu}}{c_\beta^2} \frac{\xi_{N-2}^{q'}}{\xi_N^q} \frac{\alpha_{N-2}^{q'} - C_4 e^{-\lambda u_N}}{\alpha_N^q}. \end{aligned} \tag{6.58}$$

We take care of the last factor with the help of the bound (6.48)

$$\frac{\alpha_{N-2}^{q'} - C_4 e^{-\lambda u_N}}{\alpha_N^q} \geq \frac{1}{C_2 N^2}\bigg(\frac{1}{C_2(N-2)^2} - C_5 e^{-\lambda u_N}\bigg) \geq C_5 N^{-4} \tag{6.59}$$

for N large, by choosing $\nu > \frac{2}{\lambda}$. For the second factor, we use the bound (6.49), and the Lipschitz nature of $\mathfrak{L}$ and $\widetilde{\mathbf{h}}$ on a compact set, and the fact that $|q - q'| \leq C_6 \frac{\log N}{N}$,

$$\begin{aligned} \frac{\xi_{N-2}^{q'}}{\xi_N^q} &\geq e^{-2C_3} \exp\big(N\big[\mathfrak{L}_\Lambda\big(\widetilde{\mathbf{h}}(q', 0)\big) - \widetilde{h}_0(q', 0) q'\big] \\ &\qquad - \big[\mathfrak{L}_\Lambda\big(\widetilde{\mathbf{h}}(q, 0)\big) - \widetilde{h}_0(q, 0) q\big]\big) \\ &\geq e^{-2C_3} e^{-N C_7 |q - q'|} \geq e^{-2C_3} e^{-C_7 C_6 \log N} \geq C_8 N^{-C_9}. \end{aligned} \tag{6.60}$$

Eventually, combining (6.58), (6.59) and (6.60), we obtain the lower bound, for $\mu = 4 + C_9 + \beta\nu$ and $C > 0$ a constant

$$\mathbf{P}_\beta\big(V_i > 0, 0 < i < N | NY_N = N^2 q, V_N = 0\big) \geq C N^{-\mu}. \qquad \square$$

APPENDIX A: EQUIVALENCE BETWEEN THEOREM D AND THEOREMS E AND F

Assume that Theorems E and F hold. We begin by observing that

$$\begin{aligned} d_H\bigg(\frac{S_L(l)}{\sqrt{L}}, \mathcal{S}_\beta\bigg) &\leq \frac{N_L(l)}{\sqrt{L}} d_H\bigg(\frac{S_L(l)}{N_L(l)}, \frac{\sqrt{L}}{N_L(l)} \mathcal{S}_\beta\bigg) \\ &\leq \frac{N_L(l)}{\sqrt{L}} d_H\bigg(\frac{S_L(l)}{N_L(l)}, \frac{\mathcal{S}_\beta}{a_\beta}\bigg) + \frac{N_L(l)}{\sqrt{L}} d_H\bigg(\frac{\mathcal{S}_\beta}{a_\beta}, \frac{\sqrt{L}}{N_L(l)} \mathcal{S}_\beta\bigg). \end{aligned} \tag{A.1}$$

Theorem E, and the inequality $d_H(\frac{\mathcal{S}_\beta}{a_\beta}, \frac{\sqrt{L}}{N_L(l)}\mathcal{S}_\beta) \le C|\frac{\sqrt{L}}{N_L(l)} - \frac{1}{a_\beta}|$ (C is the radius of a ball containing $\mathcal{S}_\beta$) ensure that the second term in the RHS of (A.1) converges to 0 in $P_{L,\beta}$ probability. The same convergence holds for the first term in the RHS of (A.1) and this is a consequence of Theorem F and of the inequality

$$d_H\left(\frac{S_L(l)}{N_L(l)}, \frac{\mathcal{S}_\beta}{a_\beta}\right) \le \max\left\{\left\|\widetilde{\mathcal{E}}_l^+ - \frac{\gamma_\beta^*}{2}\right\|_\infty, \left\|\widetilde{\mathcal{E}}_l^- + \frac{\gamma_\beta^*}{2}\right\|_\infty\right\} + \frac{1}{N_L(l)}.$$

Thus, Theorem D is a consequence of Theorems E and F. Using similar arguments, we can prove that Theorems E and F are implied by Theorem D but we do not give the details here.

APPENDIX B: PROOF OF LEMMA 3.2

PROOF. Since V and A_n are symmetric, we can assume that $x, x' \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and thus it is sufficient to show that the result holds for $x' = x + 1$. We will argue by induction. Since $A_0 = 0$, the $m = 0$ case is trivial. Now, we assume that the inequality holds true for $m \in \mathbb{N}$. We consider the partition function of size $m + 1$, and we can decompose it with respect to the position of V_1, that is,

$$\begin{aligned}
\mathbf{E}_{\beta,x}(e^{-\delta A_{m+1}}) &= \sum_{y\in\mathbb{Z}} \mathbf{E}_{\beta,x}(e^{-\delta(|y|+|V_2|+\cdots+|V_{m+1}|)}\mathbf{1}_{\{V_1=y\}}) \\
&= \sum_{y\in\mathbb{Z}} \mathbf{P}_\beta(U_1 = y - x)e^{-\delta|y|}\mathbf{E}_{\beta,y}(e^{-\delta A_m}) \qquad \text{(B.1)}\\
&= \sum_{y\in\mathbb{N}} R_x(y)e^{-\delta y}\mathbf{E}_{\beta,y}(e^{-\delta A_m}) + \mathbf{P}_\beta(U_1 = x)\mathbf{E}_\beta(e^{-\delta A_m}),
\end{aligned}$$

where $R_x(y) = \mathbf{P}_\beta(U_1 = y - x) + \mathbf{P}_\beta(U_1 = -y - x)$. Then we set $\bar{R}_x(y) = \sum_{y'\ge y} R_x(y')$ for $y \in \mathbb{N}$. Since $\bar{R}_x(1) + \mathbf{P}_\beta(U_1 = x) = 1$, we can rewrite the RHS in (B.1) as

$$\begin{aligned}
&\mathbf{E}_{\beta,x}(e^{-\delta A_{m+1}}) \\
&\quad = \mathbf{E}_\beta(e^{-\delta A_m}) \qquad \text{(B.2)}\\
&\qquad + \sum_{y\in\mathbb{N}} \bar{R}_x(y)[e^{-\delta y}\mathbf{E}_{\beta,y}(e^{-\delta A_m}) - e^{-\delta(y-1)}\mathbf{E}_{\beta,(y-1)}(e^{-\delta A_m})].
\end{aligned}$$

We will show that, for all $y \in \mathbb{N}$, the function $x \mapsto \bar{R}_x(y)$ is nondecreasing on $\mathbb{N}_0$. First, if $y \ge x + 1$, we obviously have

$$\bar{R}_x(y) = \sum_{y'\ge y} R_x(y') \le \sum_{y'\ge y} R_{x+1}(y') = \bar{R}_{x+1}(y). \qquad \text{(B.3)}$$

Then, if $1 \le y \le x$, since

$$\begin{aligned}
&\bar{R}_x(y) + \sum_{y'=1}^{y-1} R_x(y') + \mathbf{P}_\beta(U_1 = x) \\
&\quad = \bar{R}_{x+1}(y) + \sum_{y'=1}^{y-1} R_{x+1}(y') + \mathbf{P}_\beta(U_1 = x+1) = 1
\end{aligned} \tag{B.4}$$

and

$$\mathbf{P}_\beta(U_1 = x) + \sum_{y'=1}^{y-1} R_x(y') \ge \mathbf{P}_\beta(U_1 = x+1) + \sum_{y'=1}^{y-1} R_{x+1}(y'), \tag{B.5}$$

we immediately obtain $\bar{R}_x(y) \le \bar{R}_{x+1}(y)$. Coming back to (B.2), we use the induction hypothesis to claim that

$$e^{-\delta y}\mathbf{E}_{\beta,y}(e^{-\delta A_m}) - e^{-\delta(y-1)}\mathbf{E}_{\beta,(y-1)}(e^{-\delta A_m}) \le 0, \qquad y \in \mathbb{N}, \tag{B.6}$$

which, together with the monotonicity of $x \mapsto \bar{R}_x(y)$ yields that

$$\begin{aligned}
\mathbf{E}_{\beta,x}(e^{-\delta A_{m+1}}) &\ge \mathbf{E}_\beta(e^{-\delta A_m}) \\
&\quad + \sum_{y \in \mathbb{N}} \bar{R}_{x+1}(y)[e^{-\delta y}\mathbf{E}_{\beta,y}(e^{-\delta A_m}) - e^{-\delta(y-1)}\mathbf{E}_{\beta,(y-1)}(e^{-\delta A_m})] \\
&= \mathbf{E}_{\beta,x+1}(e^{-\delta A_{m+1}}). \qquad \square
\end{aligned}$$

Acknowledgment. We are grateful to Remco van der Hofstad for fruitful discussions.

REFERENCES

ALEXANDER, K. S. and SIDORAVICIUS, V. (2006). Pinning of polymers and interfaces by random potentials. *Ann. Appl. Probab.* **16** 636–669. MR2244428

ASMUSSEN, S. (2003). *Applied Probability and Queues*, 2nd ed. *Applications of Mathematics* (*New York*) **51**. Springer, New York. MR1978607

BRAK, R., GUTTMANN, A. J. and WHITTINGTON, S. G. (1992). A collapse transition in a directed walk model. *J. Phys. A* **25** 2437–2446. MR1164357

BRAK, R., OWCZAREK, A. L. and PRELLBERG, T. (1993). The tricritical behavior of self-interacting partially directed walks. *J. Stat. Phys.* **72** 737–772.

BRAK, R., OWCZAREK, A. L., PRELLBERG, T. and GUTTMANN, A. J. (1993). Finite-length scaling of collapsing directed walks. *Phys. Rev. E* **48** 2386–2396.

BRAK, R., DYKE, P., LEE, J., OWCZAREK, A. L., PRELLBERG, T., RECHNITZER, A. and WHITTINGTON, S. G. (2009). A self-interacting partially directed walk subject to a force. *J. Phys. A* **42** 085001, 30. MR2525477

CARAVENNA, F., DEN HOLLANDER, F. and PÉTRÉLIS, N. (2012). Lectures on random polymers. In *Probability and Statistical Physics in Two and More Dimensions*. *Clay Math. Proc.* **15** 319–393. Amer. Math. Soc., Providence, RI. MR3025394

DEMBO, A. and ZEITOUNI, O. (2010). *Large Deviations Techniques and Applications*. *Stochastic Modelling and Applied Probability* **38**. Springer, Berlin. MR2571413

DERRIDA, B., HAKIM, V. and VANNIMENUS, J. (1992). Effect of disorder on two-dimensional wetting. *J. Stat. Phys.* **66** 1189–1213. MR1156401

DERRIDA, B., GIACOMIN, G., LACOIN, H. and TONINELLI, F. L. (2009). Fractional moment bounds and disorder relevance for pinning models. *Comm. Math. Phys.* **287** 867–887. MR2486665

DOBRUSHIN, R. and HRYNIV, O. (1996). Fluctuations of shapes of large areas under paths of random walks. *Probab. Theory Related Fields* **105** 423–458. MR1402652

DOBRUSHIN, R., KOTECKÝ, R. and SHLOSMAN, S. (1992). *Wulff Construction*: *A Global Shape from Local Interaction*. *Translations of Mathematical Monographs* **104**. Amer. Math. Soc., Providence, RI. MR1181197

DURRETT, R. (2010). *Probability*: *Theory and Examples*, 4th ed. Cambridge Univ. Press, Cambridge. MR2722836

GIACOMIN, G. (2007). *Random Polymer Models*. Imperial College Press, London. MR2380992

GIACOMIN, G. (2011). *Disorder and Critical Phenomena Through Basic Probability Models*. *Lecture Notes in Math.* **2025**. Springer, Heidelberg. MR2816225

HRYNIV, O. and VELENIK, Y. (2004). Universality of critical behaviour in a class of recurrent random walks. *Probab. Theory Related Fields* **130** 222–258. MR2093763

IOFFE, D. (1994). Large deviations for the 2D Ising model: A lower bound without cluster expansions. *J. Stat. Phys.* **74** 411–432. MR1257822

IOFFE, D. (1995). Exact large deviation bounds up to T_c for the Ising model in two dimensions. *Probab. Theory Related Fields* **102** 313–330. MR1339736

IOFFE, D. and SCHONMANN, R. H. (1998). Dobrushin–Kotecký–Shlosman theorem up to the critical temperature. *Comm. Math. Phys.* **199** 117–167. MR1660207

JANSON, S. (2007). Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas. *Probab. Surv.* **4** 80–145. MR2318402

KAC, M. (1946). On the average of a certain Wiener functional and a related limit theorem in calculus of probability. *Trans. Amer. Math. Soc.* **59** 401–414. MR0016570

NGUYEN, G. B. and PÉTRÉLIS, N. (2013). A variational formula for the free energy of the partially directed polymer collapse. *J. Stat. Phys.* **151** 1099–1120. MR3063498

OWCZAREK, A. L. and PRELLBERG, T. (2007). Exact solution of semi-flexible and super-flexible interacting partially directed walks. *J. Stat. Mech. Theory Exp.* **P11010** 1–14.

SAKHANENKO, A. I. (1980). On unimprovable estimates of the rate of convergence in the invariance principle. In *Colloq. Math. Soc. János Bolyai. Nonparametric Statistical Inference* **32** 779–783. North Holland Publishing Company, Amsterdam.

SAMANTA, H. S. and THIRUMALAI, D. (2013). Exact solution of the Zwanzig–Lauritzen model of polymer crystallization under tension. *J. Chem. Phys.* **138** 104901.

SHAO, Q. M. (1995). Strong approximation theorems for independent random variables and their applications. *J. Multivariate Anal.* **52** 107–130. MR1325373

TAKÁCS, L. (1993). On the distribution of the integral of the absolute value of the Brownian motion. *Ann. Appl. Probab.* **3** 186–197. MR1202522

TORRI, N. (2015). Pinning model with heavy tailed disorder. *Stochastic Process. Appl.* Avalaible online 21 September 2015.

VAN DER HOFSTAD, R., DEN HOLLANDER, F. and KÖNIG, W. (2003). Weak interaction limits for one-dimensional random polymers. *Probab. Theory Related Fields* **125** 483–521. MR1974412

WULFF, G. (1901). Zur frage der Geschwindigkeit des Wachstums und der auflosüng der krystallflagen. *Z. Kryst. Mineral.* **34** 449.

ZWANZIG, R. and LAURITZEN, J. I. (1968). Exact calculation of the partition function for a model of two dimensional polymer crystallization by chain folding. *J. Chem. Phys.* **48** 3351.

LABORATOIRE DE MATHÉMATIQUES
JEAN LERAY UMR 6629
UNIVERSITÉ DE NANTES
2 RUE DE LA HOUSSINIÈRE
BP 92208, F-44322 NANTES CEDEX 03
FRANCE
E-MAIL: philippe.carmona@univ-nantes.fr
nicolas.petrelis@univ-nantes.fr

CENTRO DE MODELAMIENTO MATEMATICO
UNIVERSIDAD DE CHILE
AV. BEAUCHEF 851
TORRE NORTE, SANTIAGO
CHILE
E-MAIL: bao.nggia@gmail.com