A NUMERICAL APPROACH TO COPOLYMERS AT SELECTIVE INTERFACES

FRANCESCO CARAVENNA, GIAMBATTISTA GIACOMIN, AND MASSIMILIANO GUBINELLI

Abstract. We consider a model of a *random copolymer at a selective interface* which undergoes a localization/delocalization transition. In spite of the several rigorous results available for this model, the theoretical characterization of the phase transition has remained elusive and there is still no agreement about several important issues, for example the behavior of the polymer near the phase transition line. From a rigorous viewpoint non coinciding upper and lower bounds on the critical line are known.

In this paper we combine numerical computations with rigorous arguments to get to a better understanding of the phase diagram. Our main results include:

– Various numerical observations that suggest that the critical line lies *strictly* in between the two bounds.

– A rigorous statistical test based on concentration inequalities and super–additivity, for determining whether a given point of the phase diagram is in the localized phase. This is applied in particular to show that, with a very low level of error, the lower bound does not coincide with the critical line.

– An analysis of the precise asymptotic behavior of the partition function in the delocalized phase, with particular attention to the effect of rare *atypical* stretches in the disorder sequence and on whether or not in the delocalized regime the polymer path has a Brownian scaling.

– A new proof of the lower bound on the critical line. This proof relies on a characterization of the localized regime which is more appealing for interpreting the numerical data.

Keywords: Disordered Models, Copolymers, Localization Transition, Large Deviations, Corrections to Laplace estimates, Concentration of Measure, Transfer Matrix Approach, Statistical Tests

2000 MSC: 60K37, 82B44, 82B80

1. Introduction

1.1. **The model.** Let $S = \{S_n\}_{n=0,1,\dots}$ be a random walk with $S_0 = 0$ and $S_n = \sum_{j=1}^n X_j$, $\{X_j\}_j$ a sequence of IID random variables and $\mathbf{P}(X_1 = 1) = \mathbf{P}(X_1 = -1) = 1/2$. For $\lambda \geq 0$, $h \geq 0$, $N \in 2\mathbb{N}$ and $\omega = \{\omega_j\}_{j=1,2,\dots} \in \mathbb{R}^{\mathbb{N}}$ we introduce the probability measure $\mathbf{P}_{N,\omega}^{\lambda,h}$ defined by

$$\frac{\mathrm{d}\mathbf{P}_{N,\omega}^{\lambda,h}}{\mathrm{d}\mathbf{P}}(S) = \frac{1}{\widetilde{Z}_{N,\omega}^{\lambda,h}} \exp\left(\lambda \sum_{n=1}^{N} (\omega_n + h)\,\mathrm{sign}\,(S_n)\right), \tag{1.1}$$

where $\widetilde{Z}_{N,\omega}^{\lambda,h}$ is the partition function and $\mathrm{sign}\,(S_{2n})$ is set to be equal to $\mathrm{sign}\,(S_{2n-1})$ for any n such that $S_{2n} = 0$. This is a natural choice, as it is explained in the caption of Fig. 1.

For what concerns the *charges* ω we put ourselves in a *quenched set–up*: ω is a typical realization of an IID sequence of random variables (we denote by $\mathbb{P}$ its law). We suppose

Date: April 22, 2005.

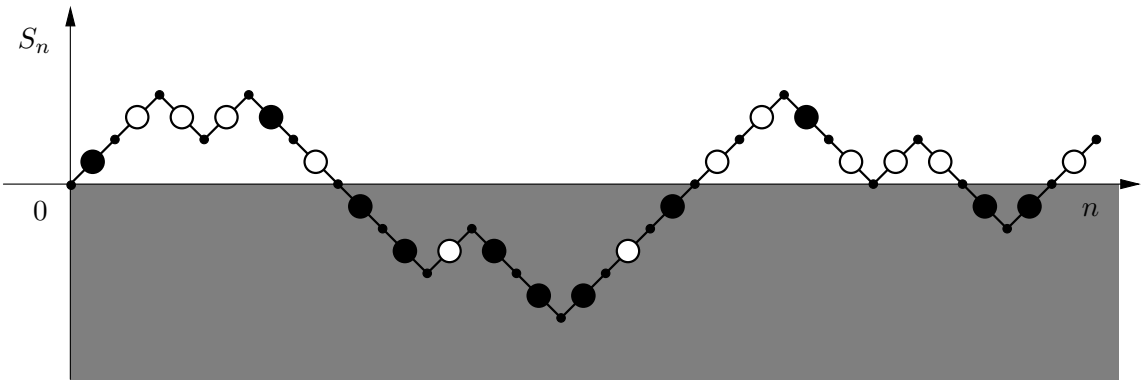

FIGURE 1. The process we have introduced is a model for a non–homogeneous polymer, or *copolymer*, near an interface, the horizontal axis, between two selective solvents, say oil (white) and water (grey). In the drawing the monomer *junctions* are the small black rounds and the monomers are the bonds of the random walk. The big round in the middle of each monomer gives the sign of the charge (white = positive charge = hydrophobic monomer, black = negative charge = hydrophilic monomer). When $h > 0$ water is the unfavorable solvent and the question is whether the polymer is *delocalized* in oil or if it is still more profitable to place a large number of monomers in the preferred solvent, leading in such a way to the *localization at the interface* phenomenon. The conventional choice of sign(0) we have made reflects the fact that the charge is assigned to bonds rather than points.

that

$$\mathsf{M}(\alpha) := \mathbb{E}\left[\exp\left(\alpha\omega_1\right)\right] < \infty\,, \tag{1.2}$$

for every α and that $\mathbb{E}\left[\omega_1\right] = 0$. Moreover we fix $\mathbb{E}[{\omega_1}^2] = 1$.

1.2. **The free energy and the phase diagram.** We introduce the free energy of the system

$$f(\lambda, h) \;=\; \lim_{N\to\infty} \frac{1}{N} \log \widetilde{Z}^{\lambda,h}_{N,\omega}. \tag{1.3}$$

The limit has to be understood in the $\mathbb{P}\left(\,\mathrm{d}\omega\right)$–almost sure sense, or in the $\mathbb{L}_1\left(\mathbb{P}\right)$ sense, and $f(\lambda, h)$ does not depend on ω. A proof of the existence of such a limit goes along a standard superadditive argument and we refer to [14] for the details, see however § 2.1 below. By convexity arguments one easily sees that the free energy is a continuous function.

We observe that

$$f(\lambda, h) \;\geq\; \lambda h. \tag{1.4}$$

In fact if we set $\Omega_N^+ = \{S :\, S_n > 0 \text{ for } n = 1, 2, \ldots, N\}$

$$\begin{aligned}\frac{1}{N} \log \widetilde{Z}^{\lambda,h}_{N,\omega} \geq \frac{1}{N} \log \mathbf{E}\left[\exp\left(\lambda \sum_{n=1}^{N} \left(\omega_n + h\right) \operatorname{sign}\left(S_n\right)\right); \Omega_N^+\right] \\ = \frac{\lambda}{N} \sum_{n=1}^{N} \left(\omega_n + h\right) \;+\; \frac{1}{N} \log \mathbf{P}\left(\Omega_N^+\right) \;\overset{N\to\infty}{\longrightarrow}\; \lambda h,\end{aligned} \tag{1.5}$$

where the limit has to be understood in the $\mathbb{P}(\,\mathrm{d}\omega)$–almost sure sense: notice that we have used the law of large numbers. We have of course also applied the well known fact that

$\mathbf{P}\left(\Omega_N^+\right)$ behaves like $N^{-1/2}$ for N large [12, Ch. III]. In view of (1.3) and of (1.5) we partition the phase diagram in the following way:

- The localized region: $\mathcal{L} = \{(\lambda, h) : f(\lambda, h) - \lambda h > 0\}$;
- The delocalized region: $\mathcal{D} = \{(\lambda, h) : f(\lambda, h) - \lambda h = 0\}$.

This phase diagram decomposition does correspond to different behaviors of the trajectories of the copolymer: we will come back to this important issue in § 1.4.

We sum up in the following theorem what is known about the phase diagram of the model.

Theorem 1.1. *There exists a continuous increasing function $h_c : [0, \infty) \longrightarrow [0, \infty)$, $h_c(0) = 0$, such that*

$$\mathcal{L} = \{(\lambda, h) : h < h_c(\lambda)\} \quad \text{and} \quad \mathcal{D} = \{(\lambda, h) : h \geq h_c(\lambda)\}. \tag{1.6}$$

Moreover

$$\underline{h}(\lambda) := \frac{1}{4\lambda/3} \log \mathsf{M}(-4\lambda/3) \leq h_c(\lambda) \leq \frac{1}{2\lambda} \log \mathsf{M}(-2\lambda) =: \overline{h}(\lambda). \tag{1.7}$$

This implies that the slope at the origin belongs to $[2/3, 1]$, in the sense that the inferior limit of $h_c(\lambda)/\lambda$ as $\lambda \searrow 0$ is not smaller than $2/3$ and the superior limit is not larger than 1.

Remark 1.2. In [15] it is proven that the limit of $h_c(\lambda)/\lambda$ as $\lambda \searrow 0$ does exist and it is independent of the distribution of ω_1, at least when ω_1 is a bounded symmetric random variable or when ω_1 is a standard Gaussian variable [15]. This *universal character* of the slope at the origin makes this quantity very interesting.

Theorem 1.1 is a mild generalization of the results proven in [3] and [4]: the extension lies in the fact that ω_1 is not necessarily symmetric and a proof of it requires minimal changes with respect to the arguments in [4]. The lower bound in (1.7) is actually proven explicitly in Appendix B (see also Section 3), but we stress that we present this proof because it is a new one and because it gives some insight on the computational results. For what follows we set

$$h^{(m)}(\lambda) = \frac{1}{2m\lambda} \log \mathsf{M}(-2m\lambda), \tag{1.8}$$

for $m > 0$. Observe that the curves $\underline{h}(\cdot)$ and $\overline{h}(\cdot)$ defined in (1.7) correspond respectively to $m = 2/3$ and $m = 1$, and that $\frac{\mathrm{d}}{\mathrm{d}\lambda} h^{(m)}(\lambda)|_{\lambda=0} = m$.

Remark 1.3. Notice that one can write

$$\frac{\mathrm{d}\mathbf{P}_{N,\omega}^{\lambda,h}}{\mathrm{d}\mathbf{P}}(S) = \frac{1}{Z_{N,\omega}^{\lambda,h}} \exp\left(-2\lambda \sum_{n=1}^{N} (\omega_n + h)\, \Delta_n\right), \tag{1.9}$$

with $\Delta_n = (1 - \mathrm{sign}(S_n))/2$ and $Z_{N,\omega} := Z_{N,\omega}^{\lambda,h}$ a new partition function which coincides with $\widetilde{Z}_{N,\omega} \exp\left(-\lambda \sum_{n=1}^{N} (\omega_n + h)\right)$ and therefore we have

$$\mathrm{F}(\lambda, h) := \lim_{N \to \infty} \frac{1}{N} \log Z_{N,\omega} = f(\lambda, h) - \lambda h. \tag{1.10}$$

This limit of course has to be interpreted in the $\mathbb{P}(\mathrm{d}\omega)$–a.s. sense. We stress that, even if equivalent to $\widetilde{Z}_{N,\omega} \exp(-\lambda h N)$ in the Laplace asymptotic sense, $Z_{N,\omega}$ turns out to be

substantially more useful. This had been already realized in [3], but for our results looking at $Z_{N,\omega}$, rather than $\widetilde{Z}_{N,\omega}$, is even more essential. Moreover from now on $\mathrm{F}(\lambda,h)$, rather than $f(\lambda,h)$, will be for us the free energy.

We will use repeatedly also the partition function associated to the model *pinned* at the right endpoint:

$$Z^{\lambda,h}_{N,\omega}(x) \,:=\, \mathbf{E}\left[\exp\left(-2\lambda\sum_{n=1}^{N}(\omega_n+h)\,\Delta_n\right);\, S_N=x\right]. \tag{1.11}$$

It is worth recalling that one can substitute $Z^{\lambda,h}_{N,\omega}$ with $Z^{\lambda,h}_{N,\omega}(x)$, any fixed even x, in (1.10) and the limit is unchanged, see e.g. [3] or [14].

1.3. **A random walk excursions viewpoint.** We present here a different viewpoint on the process: this turns out to be useful for the intuition and it will be used in some technical steps.

We call η the first return time of the walk S to 0, that is $\eta := \inf\{n\geq 1 : S_n = 0\}$, and set $K(2n) := \mathbf{P}(\eta = 2n)$ for $n\in\mathbb{N}$. It is well known that $K(\cdot)$ is decreasing on the even natural numbers and

$$\lim_{x\in 2\mathbb{N}, x\to\infty} x^{3/2}K(x) = \sqrt{2/\pi}, \tag{1.12}$$

see e.g. [12, Ch. 3]. Let the IID sequence $\{\eta_j\}_{j=1,2,\ldots}$ denote the inter–arrival times at 0 for S, and we set $\tau_k := \eta_0+\ldots+\eta_k$. If we introduce also $\ell_N = \max\{j\in\mathbb{N}\cup\{0\} : \tau_j\leq N\}$, then by exploiting the up–down symmetry of the excursions of S we directly obtain

$$\begin{aligned} Z_{N,\omega}(0) \,&=\, \mathbf{E}\left[\prod_{j=1}^{\ell_N}\varphi\left(\lambda\sum_{n=\tau_{j-1}+1}^{\tau_j}\omega_n+\lambda h\eta_j\right);\tau_{\ell_N}=N\right] \\ &= \sum_{l=0}^{N}\sum_{\substack{x_0,\ldots,x_l\in 2\mathbb{N}\\ 0=:x_0<\ldots<x_l:=N}}\prod_{i=1}^{l}\varphi\left(\lambda\sum_{n=x_{i-1}+1}^{x_i}\omega_n+\lambda h(x_i-x_{i-1})\right)K(x_i-x_{i-1}), \end{aligned} \tag{1.13}$$

with $\varphi(t) := (1+\exp(-2t))/2$. Of course the formula for $Z_{N,\omega}$ is just slightly different.

Formula (1.13) reflects the fact that what really matters for the copolymer are the return times to the interface.

1.4. **Known and conjectured path properties.** The question of whether splitting the phase diagram into the regions $\mathcal{L}$ and $\mathcal{D}$ does correspond to really different path behaviors has a positive answer, at least if we do not consider the critical case, that is if we consider the path behavior for $(\lambda,h)\in\mathcal{L}$ and for (λ,h) in the interior of $\mathcal{D}$. However, while the localized regime is rather well understood, the delocalized one remains somewhat elusive (we take up this point again in Section 4.1). More precisely:

- For $(\lambda,h)\in\mathcal{L}$ one knows that the polymer is going to stay very close to the interface, *essentially* at distance $O(1)$ and the polymer becomes finitely recurrent for $N\to\infty$. Due to the disordered distribution of the charges, even the most elementary results in this direction require a careful formulation and we prefer to refer to [1], [2] and [24].
- For (λ,h) in the interior of $\mathcal{D}$ one can prove by large deviation arguments that there are $o(N)$ visits to the unfavorable solvent and by more sophisticate arguments that these visits are actually $O(\log N)$ [15]. These results are in sharp contrast

with what happens in $\mathcal{L}$ and in this sense they are satisfactory. However they give at the same time still a weak information on the paths, above all if compared to what is available for non disordered models, see e.g. [21], [11], [8] and references therein, namely Brownian scaling, which in turn is a consequence of the fact that all the visits in the unfavorable solvent happen very close to the boundary points, that is the origin, under the measure $\mathbf{P}_{N,\omega}^{\lambda,h}$. In non disordered models one can in fact prove that the polymer becomes transient and that it visits the unfavorable solvent, or any point below a fixed level, only a finite number of times. Recently it has been shown [15] that such a result cannot hold as stated, at least for $h < \overline{h}(\lambda)$, for the disordered copolymer. However the results in [15] leave open the possibility of Brownian scaling in the whole delocalized region.

1.5. **Outline of the results.** Formula (1.7) leaves an important gap, that hides the only partial understanding of the nature of this delocalization/localization transition. Our purpose is to go toward filling this gap: our results are both of theoretical and numerical nature. At the same time we address the delocalization issues raised in § 1.4, which are intimately related with the precise asymptotic behavior of $Z_{N,\omega}$ and of $Z_{N,\omega}(0)$. More precisely:

(1) In Section 2 we present a statistical test with explicit error bounds, see Proposition 2.2, based on super–additivity and concentration inequalities, to state that a point (λ, h) is localized. We apply this test to show that, with a very low level of error, the lower bound $h = \underline{h}(\lambda)$ defined in (1.7) does not coincide with the critical line.
(2) In Section 3 we give the outline of a new proof of the main result in [4]. The details of the proof are in Appendix B and we point out in particular Proposition B.2, that gives a necessary and sufficient condition for localization. This viewpoint on the transition, derived from [15, Section 4], helps substantially in interpreting the *irregularities* in the behavior of $\{Z_{N,\omega}\}_N$ as $N \nearrow \infty$.
(3) In Section 4 we pick up the conjecture of Brownian scaling in the delocalized regime both in the intent of testing it and in trying to asses with reasonable confidence that (λ, h) is in the interior of $\mathcal{D}$. In particular, we present quantitative evidences in favor of the fact that the upper bound $h = \overline{h}(\lambda)$ defined in (1.7) is strictly greater than the critical line. We stress that this is a very delicate issue, since delocalization, unlike localization, does not appear to be reducible to a finite volume issue.
(4) Finally, in Section 5, we report the results of a numerical attempt to determine the critical curve. While this issue has to be treated with care, mostly for the reasons raised in point 4 above, we observe a surprising phenomenon: the critical curve appears to be very close to $h^{(m)}(\cdot)$ for a suitable value of m. By the universality result proven in [15], building on the free energy Brownian scaling result proven in [3], the slope at the origin of $h_c(\cdot)$ does not depend on the law of ω. Therefore if really $h^{(m)}(\cdot) = h_c(\cdot)$, since the slope at the origin of $h^{(m)}(\cdot)$ is m, m is the universal constant we are looking for. We do not believe that the numerical evidence allows to make a clear cut statement, but what we observe is compatible with such a possibility.

We point out that our numerical results are based on a numerical computation of the partition function $Z_{N,\omega}$, exploiting the standard transfer–matrix approach (this item is discussed in more details in Appendix A).

1.6. **A quick overview of the literature.** The *copolymer in the proximity of an interface* problem has a long literature, but possibly the first article that attracted the attention of mathematicians is [16]. Here we are going to focus on very specific issues and the most interesting for our purposes is that in the physical literature both the conjecture that $\underline{h}(\cdot) = h_c(\cdot)$ (cf. [20] and [26]) and that $\overline{h}(\cdot) = h_c(\cdot)$ (cf. [27]) are set forth. The approaches are non rigorous, mostly based on replica computations, with the exception of [20] whose method is the real space renormalization technique for one–dimensional disordered systems first proposed in [13] in the context of quantum Ising model with transverse field and then applied with remarkably precise results to random walk in random environment, see e. g. [17]. The result in [4], that $\underline{h}(\cdot) \le h_c(\cdot)$, is obtained by exploiting the path behavior of the copolymer near criticality suggested in [20]. This strategy may by summed up by: the localized polymer close to criticality is mostly delocalized in the upper half–plane and it keeps in the lower half–plane only the rare portions with an atypically negative charge. The numerical results that we set forth in this work are saying that this strategy is not good enough.

At the opposite end, the result $\overline{h}(\cdot) \ge h_c(\cdot)$, albeit relatively subtle, is absolutely elementary to prove [3]. And such a bound does not depend at all on the details of the walk: any non trivial null recurrent walk with increments in $\{-1, 0, +1\}$ leads to the same upper bound. This suggests that such a bound is too rough. One can however prove that the standard procedure for obtaining upper bounds that goes under the name of *constrained annealing* cannot improve such a bound [7]. This is in any case far from being a proof that $\overline{h}(\cdot) = h_c(\cdot)$, and in fact the numerics suggest that this is not the case.

In the literature one finds also a large number of numerical works on copolymers, we mention here for example [9], [25] and references therein. As far as we have seen, the attention is often shifted toward different aspects, notably of course the issue of critical exponents, and the more complex model in which the polymer is not directed but rather self–avoiding, see [9] and [25] also for some rigorous results and references in such a direction.

Our work has been led rather by the idea that understanding the precise location of the critical curve is a measure of our understanding of the nature of the transition. Understanding that, in turn, could promote an advance on the mathematical analysis of the copolymer and, more generally, of this kind of disordered models.

2. A statistical test for the localized phase

2.1. **Checking localization at finite volume.** At an intuitive level one is led to believe that, when the copolymer is localized, it should be possible to detect it by looking at the system before the infinite volume limit. This intuition is due to the fact that in the localized phase the length of each excursion is finite, therefore for N *much larger* that the *typical* excursion length one should already observe the localization phenomenon in a quantitative way. The system being disordered of course does not help, because it is more delicate to make sense of what *typicality* means in a non translation invariant set–up. However the translation invariance can be recovered by averaging and in fact it turns out to be rather easy to give a precise meaning to the intuitive idea we have just mentioned. The key word here is super–additivity of the averaged free energy.

In fact by considering only the S trajectories such that $S_{2N} = 0$ and by applying the Markov property of S one directly verifies that for any $N, M \in \mathbb{N}$

$$Z_{2N+2M,\omega}(0) \;\geq\; Z_{2N,\omega}(0)\, Z_{2M,\theta^{2N}\omega}(0), \tag{2.1}$$

$(\theta\omega)_n = \omega_{n+1}$, and therefore

$$\{\, \mathbb{E} \log Z_{2N,\omega}(0)\}_{N=1,2,\ldots} \tag{2.2}$$

is a super–additive sequence, which immediately entails the existence of the limit of $\mathbb{E}[\log Z_{2N,\omega}(0)]/2N$ and the fact that this limit coincides with the supremum of the sequence. Therefore from the existence of the quenched free energy we have that

$$\mathrm{F}(\lambda, h) \;=\; \sup_N \frac{1}{2N} \mathbb{E} \log Z_{2N,\omega}(0)\,. \tag{2.3}$$

In a more suggestive way one may say that:

$$(\lambda, h) \in \mathcal{L} \quad \Longleftrightarrow \quad \text{there exists } N \in \mathbb{N} \text{ such that } \mathbb{E} \log Z_{2N,\omega}(0) > 0\,. \tag{2.4}$$

The price one pays for working with a disordered system is precisely in taking the $\mathbb{P}$–expectation and from the numerical viewpoint it is an heavy price: even with the most positive attitude one cannot expect to have access to $\mathbb{E} \log Z_{2N,\omega}(0)$ by direct numerical computation for N above 10. Of course in principle small values of N may suffice (and they do in some cases, see Remark 2.1), but they do not suffice to tackle the specific issue we are interested in. We elaborate at length on this interesting issue in § 2.4.

Remark 2.1. An elementary application of the localization criterion (2.4) is obtained for $N = 1$: $(\lambda, h) \in \mathcal{L}$ if

$$\mathbb{E}\left[\log\left(\frac{1}{2} + \frac{1}{2}\exp\left(-2\lambda\left(\omega_1 + \omega_2 + 2h\right)\right)\right)\right] > 0. \tag{2.5}$$

In the case $\mathbb{P}(\omega_1 = \pm 1) = 1/2$ from (2.5) we obtain that for λ sufficiently large $h_c(\lambda) > 1 - c/\lambda$, with $c = (1/4)\log(2\exp(4) - 1) \approx 1.17$. From $\underline{h}(\cdot)$ we obtain the same type of bound, with $c = (3/4)\log 2 \approx 0.52$. This may raise some hope that for λ large an explicit, possibly computer assisted, computation for small values of N of $\mathbb{E} \log Z_{2N,\omega}(0)$ could lead to new estimates. This is not the case, as we show in § 2.4.

2.2. **Testing by using concentration.** In order to decide whether $\mathbb{E} \log Z_{2N,\omega}(0) > 0$ we resort to a Montecarlo evaluation of $\mathbb{E} \log Z_{2N,\omega}(0)$ that can be cast into a statistical test with explicit error bound by means of concentration of measure ideas. This procedure is absolutely general, but we have to choose a set–up for the computations and we take the simplest: $\mathbb{P}(\omega_1 = +1) = \mathbb{P}(\omega_1 = -1) = 1/2$. The reason for this choice is twofold:

- if ω_1 is a bounded random variable, a Gaussian concentration inequality holds and if ω is symmetric and it takes only two values then one can improve on the explicit constant in such an inequality. This speeds up in a non negligible way the computations;
- generating *true randomness* is out of reach, but playing head and tail is certainly the most elementary case in such a far reaching task (the random numbers issue is briefly discussed in Appendix A too).

A third reason to restrict testing to the Bernoulli case is explained at the end of the caption of Table 2.

We start the testing procedure by stating the null hypothesis:

$$\text{H0}: \quad \mathbb{E}\log Z_{2N,\omega}(0) \le 0. \tag{2.6}$$

N in H0 can be chosen arbitrarily. We stress that refusing H0 implies $\mathbb{E}\log Z_{2N,\omega}(0) > 0$, which by (2.4) implies localization.

The following concentration inequality for Lipschitz functions holds for the uniform measure on $\{-1,+1\}^N$: for every function $G_N : \{-1,+1\}^N \to \mathbb{R}$ such that $|G_N(\omega) - G_N(\omega')| \le C_{\text{Lip}}\sqrt{(\sum_{n=1}^N(\omega_n - \omega'_n)^2)}$, where C_{Lip} a positive constant and $G_N(\omega)$ is an abuse of notation for $G_N(\omega_1,\dots,\omega_N)$, one has

$$\mathbb{E}\left[\exp\left(\alpha\left(G_N(\omega) - \mathbb{E}[G_N(\omega)]\right)\right)\right] \le \exp\left(\alpha^2 C_{\text{Lip}}^2\right), \tag{2.7}$$

for every α. Inequality (2.7) with an extra factor 4 at the exponent can be extracted from the proof of Theorem 5.9, page 100 in [18]. Such an inequality holds for variables taking values in $[-1,1]$: the factor 4 can be removed for the particular case we are considering (see [18, p. 110–111]). In our case $G_N(\omega) = \log Z_{2N,\omega}(0)$. By applying the Cauchy–Schwarz inequality one obtains that G_N is Lipschitz with $C_{\text{Lip}} = 2\lambda\sqrt{N}$. Let us now consider an IID sequence $\{G_N^{(i)}(\omega)\}_i$ with $G_N^{(1)}(\omega) = G_N(\omega)$: if H0 holds then we have that for every $n \in \mathbb{N}$, $u > 0$ and $\alpha = un/8\lambda^2 N$

$$\begin{aligned}
\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n G_N^{(i)}(\omega) \ge u\right) &\le \mathbb{E}\left[\exp\left(\frac{\alpha}{n}\left(G_N(\omega) - \mathbb{E}[G_N(\omega)]\right)\right)\right]^n \exp\left(-\alpha\left(u - \mathbb{E}[G_N(\omega)]\right)\right) \\
&\le \exp\left(\frac{4\alpha^2\lambda^2 N}{n} - \alpha u\right) \\
&= \exp\left(-\frac{u^2 n}{16\lambda^2 N}\right).
\end{aligned} \tag{2.8}$$

Let us sum up what we have obtained:

Proposition 2.2. *Let us call $\widehat{u}_n$ the average of a sample of n independent realizations of $\log Z_{2N,\omega}^{\lambda,h}(0)$. If $\widehat{u}_n > 0$ then we may refuse H0, and therefore $(\lambda,h) \in \mathcal{L}$, with a level of error not larger than $\exp\left(-\widehat{u}_n^2 n/16\lambda^2 N\right)$.*

2.3. **Numerical tests.** We report in Table 1 the most straightforward application of Proposition 2.2, obtained by a numerical computation of $\log Z_N$ for a sample of n independent environments ω. We aim at seeing how far above $\underline{h}(\cdot)$ one can go and still claim localization, keeping a reasonably small probability of error.

Remark 2.3. One might be tempted to interpolate between the values in Table 1, or possibly to get results for small values of λ in order to extend the result of the test to the slope of the critical curve in the origin. However the fact that $h_c(\lambda)$ is strictly increasing does not help much in this direction and the same is true for the finer result, proven in [5], that $h_c(\lambda)$ can be written as $U(\lambda)/\lambda$, $U(\cdot)$ a convex function.

λ	0.3	0.6	1
h	0.22	0.41	0.58
p–value	1.5×10^{-6}	9.5×10^{-3}	1.6×10^{-5}
$\underline{h}(\lambda)$	0.195	0.363	0.530
$\overline{h}(\lambda)$	0.286	0.495	0.662
N	300000	500000	160000
n	225000	330000	970000
C. I. 99%	7.179 ± 0.050	9.011 ± 0.045	7.643 ± 0.025

TABLE 1. According to our numerical computations, the three pairs (λ, h) are in $\mathcal{L}$ and this has been tested with the stated p–values (or probability/level of error). We report the values of $\overline{h}(\lambda)$ and $\underline{h}(\lambda)$ for reference. Of course in these tests there is quite a bit of freedom in the choice of n and N: notice that N enters in the evaluation of the p–value also because a larger value of N yields a larger value of $\mathbb{E} \log Z^{\lambda,h}_{2N,\omega}(0)$. In the last line we report standard Gaussian 99% confidence intervals for $\mathbb{E} \log Z^{\lambda,h}_{2N,\omega}(0)$. Of course the p–value under the Gaussian assumption turns out to be totally negligible.

2.4. **Improving on $\underline{h}(\cdot)$ is uniformly hard.** One can get much smaller p–values at little computational cost by choosing h *just above* $\underline{h}(\lambda)$. As a matter of fact a natural choice is for example $h = h^{(0.67)}(\lambda) > \underline{h}(\lambda)$, recall (1.8), for a set of values of λ, and this is part of the content of Table 2: in particular $\mathbb{E} \log Z^{\lambda, h^{(0.67)}(\lambda)}_{2N_+,\omega}(0) > 0$ with a probability of error smaller than 10^{-5} for the values of λ between 0.1 and 1. However we stress that for some of these λ's we have a much smaller p–value, see the caption of Table 2, and that the content of this table is much richer and it approaches also the question of whether or not a symbolic computation or some other form of computer assisted argument could lead to $h_c(\lambda) > \underline{h}(\lambda)$ for some λ, and therefore for λ in an interval. Since such an argument would require N to be *small*, intuitively the hope resides in large values of λ, recall also Remark 2.1. It turns out that one needs in any case N larger than 700 in order to observe a localization phenomenon at $h^{(0.67)}(\lambda)$. We now give some details on the procedure that leads to Table 2.

λ	$0.05(\star)$	0.1	0.2	0.4	0.6	1	$2(\star)$	$4(\star\star)$	$8(\star\star)$
N_+	750000	190000	40000	9500	4250	1800	900	800	800
N_-	600000	130000	33000	7500	3650	1550	750	700	700

TABLE 2. For a given λ, both $\mathbb{E} \log Z^{\lambda, h^{(0.67)}(\lambda)}_{2N_+,\omega}(0) > 0$ and $\mathbb{E} \log Z^{\lambda, h^{(0.67)}(\lambda)}_{2N_-,\omega}(0) < 0$ with a probability of error smaller than 10^{-5} (and in some cases much smaller than that). Instead for the two cases marked by a $(\star)$ the level of error is rather between 10^{-2} and 10^{-3}. For large values of λ, the two cases marked with $(\star\star)$, it becomes computationally expensive to reach small p–values. However, above $\lambda = 3$ one observes that the values of $Z_{2N,\omega}(0)$ essentially do not depend anymore on the value of λ. This can be interpreted in terms of convergence to a limit ($\lambda \to \infty$) model, as it is explained in Remark 2.4. If we then make the hypothesis that this limit model sharply describes the copolymer along the curve $(\lambda, h^{(m)}(\lambda))$ for λ sufficiently large and we apply the concentration inequality, then the given values of N_+ and N_- are tested with a very small probability of error. Since the details of such a procedure are quite lengthy we do not report them here. We have constructed (partial) tables also for different laws of ω, notably $\omega_1 \sim N(0,1)$, and they turned out to yield larger, at times substantially larger, values of $N_\pm(\lambda)$.

First and foremost, the concentration argument that leads to Proposition 2.2 is symmetric and it works for deviations below the mean as well as above. So we can, in the very same way, test the null hypothesis $\mathbb{E}\log Z_{2N,\omega}(0) > 0$ and, possibly, refuse it if $\hat{u}_n < 0$, exactly with the same p–value as in Proposition 2.2. Of course an important part of Proposition 2.2 was coming from the finite volume localization condition (2.4): we do not have an analogous statement for delocalization (and we do not expect that there exists one). But, even if $\mathbb{E}\log Z_{2N,\omega}(0) \le 0$ does not imply delocalization, it says at least that it is pointless to try to prove localization by looking at a system of that size.

In Table 2 we show two values of the system size N, N_+ and N_-, for which, at a given λ, one has that $\mathbb{E}\log Z_{2N_+,\omega}(0) > 0$ and $\mathbb{E}\log Z_{2N_-,\omega}(0) < 0$ with a fixed probability of error (specified in the caption of the Table). It is then reasonable to guess that the transition from negative to positive values of $\mathbb{E}\log Z_{\cdot,\omega}(0)$ happens for $N \in (N_-, N_+)$. There is no reason whastoever to expect that $\mathbb{E}\log Z_{N,\omega}(0)$ should be monotonic in N but according to our numerical result it is not unreasonable to expect that monotonicity should set in for N large or, at least, that for $N < N_-$ (respectively $N > N_+$) $\mathbb{E}\log Z_{2N,\omega}(0)$ is definitely negative (respectively positive).

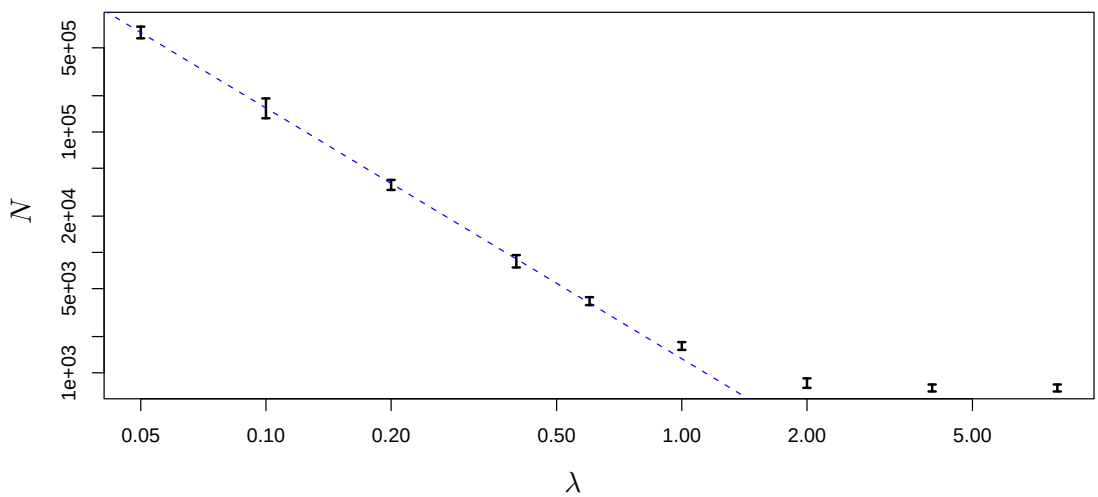

FIGURE 2. A graphical representation of Table 2. The plot is log–log, and a λ^{-c} behavior is rather evident, c is about 2.08. This can be nicely interpreted in terms of the coarse graining technique in the proof of the weak interaction scaling limit of the free energy in [3]: from that argument one extracts that if λ is small the excursions that give a contribution to the free energy have *typical* length λ^{-2} and that in the limit the polymer is just made up by this type of excursions. One therefore expects that it suffices a system of size $N(\lambda)$, with $\lim_{\lambda\searrow 0}\lambda^2 N(\lambda) = +\infty$, to observe localization if $m < h_c'(0)$, $h = h^{(m)}(\lambda) = m\lambda(1+o(1))$ and λ is small.

Remark 2.4. As pointed out in the caption of Table 2, from numerics one observes a very sharp convergence to a λ independent behavior as λ becomes large, along the line $h = h^{(m)}(\lambda)$. This is easily interpreted if one observes that $h^{(m)}(\lambda) = 1 - ((\log 2)/2m\lambda) +$

$O(\exp(-4m\lambda))$ so that

$$\lim_{\lambda\to\infty} \exp\left(-2\lambda \sum_{n=1}^{N} (\omega_n + h)\,\Delta_n\right) = \exp\left(\frac{\log 2}{m}\sum_{n=1}^{N}\Delta_n\right)\mathbf{1}_{\{\sum_{n=1}^{N}\Delta_n(1+\omega_n)=0\}}(S). \quad (2.9)$$

This corresponds to the model where a positive charge never enters the lower half-plane and where the energy of a configuration is proportional to the number of negative charges in the lower half-plane.

3. Lower bound strategies versus the true strategy

3.1. **An approach to lower bounds on the critical curve.** In this section we give an outline of a new derivation of the lower bound

$$\underline{h}(\lambda) \le h_c(\lambda), \quad (3.1)$$

with $\underline{h}(\lambda)$ defined in (1.7). The complete proof may be found in Appendix B. The argument takes inspiration from the ideas used in the proof of Proposition 3.1 in [15] and, even if it is essentially the proof of [5] in disguise, in the sense that the selection of the random walk trajectories that are kept and whose energy contribution is evaluated does not differ too much (in a word: the *strategy* of the polymer is similar), it is however conceptually somewhat different and it will naturally lead to some considerations on the precise asymptotic behavior of $Z_{N,\omega}$ in the delocalized phase and even in the localized phase close to criticality.

The first step in our proof of (3.1) is a different way of looking at localization. For any fixed positive number C we introduce the stopping time (with respect to the natural filtration of the sequence $\{\omega_n\}$) $T^C = T^{C,\lambda,h}(\omega)$ defined by

$$T^{C,\lambda,h}(\omega) := \inf\{N \in 2\mathbb{N} :\ Z^{\lambda,h}_{N,\omega}(0) \ge C\}\,. \quad (3.2)$$

The key observation is that if $\mathbb{E}[T^C] < \infty$ for some $C > 1$, then the polymer is localized. Let us sketch a proof of this fact (for the details, see Proposition B.2): notice that by the very definition of T^C we have $Z_{T^C(\omega),\omega}(0) \ge C$. Now the polymer that is in zero at $T^C(\omega)$ is equivalent to the original polymer, with a translated environment $\omega' = \theta^{T^C}\omega$, and setting $T_2(\omega) := T^C(\omega')$ we easily get $Z_{T_1(\omega)+T_2(\omega),\omega}(0) \ge C^2$ (we have put $T_1(\omega) := T^C(\omega)$). Notice that the new environment ω' is still typical, since T^C is a stopping time, so that T_2 is independent of T_1 and has the same law. This procedure can be clearly iterated, yielding an IID sequence $\{T_i(\omega)\}_{i=1,2,\dots}$ that gives the following lower bound on the partition function:

$$Z_{T_1(\omega)+\ldots+T_n(\omega),\omega}(0) \ge C^n\,. \quad (3.3)$$

From this bound one easily obtains that

$$\mathrm{F}(\lambda,h) \overset{\text{a.s.}}{=} \lim_{n\to\infty} \frac{\log Z_{T_1(\omega)+\ldots+T_n(\omega),\omega}(0)}{T_1(\omega)+\ldots+T_n(\omega)} \ge \frac{\log C}{\mathbb{E}[T^C]}\,, \quad (3.4)$$

where we have applied the strong law of large numbers, and localization follows since by hypothesis $C > 1$ and $\mathbb{E}[T^C] < \infty$.

Remark 3.1. It turns out that also the reciprocal of the claim just proved holds true, that is *the polymer is localized if and only if* $\mathbb{E}[T^C] < \infty$, with an arbitrary choice of $C > 1$, see Proposition B.2. In fact the case $\mathbb{E}[T^C] = \infty$ may arise in two different ways:

(1) the variable T^C is defective, $\mathbb{P}[T^C = \infty] > 0$: in this case with positive probability $\{Z_{N,\omega}(0)\}_N$ is a bounded sequence, and delocalization follows immediately;
(2) the variable T^C is proper with infinite mean, $\mathbb{P}[T^C = \infty] = 0$, $\mathbb{E}[T^C] = \infty$: in this case we can still build a sequence $\{T_i(\omega)\}_{i=1,2,\dots}$ defined as above and this time the lower bound (3.3) has *subexponential* growth. Moreover it can be shown that in this case the lower bound (3.3) gives the true free energy, cf. Lemma B.1, which therefore is zero, so that delocalization follows also in this case.

As a matter of fact, it is highly probable that in the interior of the delocalized phase $Z_{N,\omega}(0)$ vanishes $\mathbb{P}(\,\mathrm{d}\omega)$–a.s. when $N \to \infty$ and this would rule out the scenario (2) above, saying that for $C > 1$ the random variable T^C must be either integrable or defective. We take up again this point in Sections 4 and 5: we feel that this issue is quite crucial in order to fully understand the delocalized phase of disordered models.

Remark 3.2. Dealing directly with T^C may be difficult. Notice however that if one finds a random time (by this we mean simply an integer–valued random variable) $T = T(\omega)$ such that

$$Z_{T(\omega),\omega}(0) \geq C > 1\,, \qquad \text{with } \mathbb{E}[T] < \infty\,, \tag{3.5}$$

then localization follows. This is simply because this implies $T^C \leq T$ and hence $\mathbb{E}[T^C] < \infty$. Therefore localization is equivalent to the condition $\log Z_{T(\omega),\omega}(0) > 0$ for an *integrable* random time $T(\omega)$: we would like to stress the analogy between this and the criterion for localization given in § 2.1, see (2.4).

Now we can turn to the core of our proof: we are going to show that for every (λ, h) with $h < \underline{h}(\lambda)$ we can build a random time $T = T(\omega)$ that satisfies (3.5). The construction of T is based on the idea that for $h > 0$ if localization prevails is because of rare ω–stretches that invite the polymer to spend time in the lower half–plane in spite of the action of h.

The strategy we use consists in looking for q–atypical stretches of length at least $M \in 2\mathbb{N}$, where $q < -h$ is the average charge of the stretch. Rephrased a bit more precisely, we are looking for the smallest $n \in 2\mathbb{N}$ such that $\sum_{i=n-k+1}^{n} \omega_i/k < q$ for some even integer $k \geq M$. It is well known that such a random variable grows, in the sense of Laplace, as $\exp(\Sigma(q)M)$ for $M \to \infty$, where $\Sigma(q)$ is the Cramer functional

$$\Sigma(q) := \sup_{\alpha \in \mathbb{R}} \{\alpha q - \log \mathsf{M}(\alpha)\}\,. \tag{3.6}$$

One can also show without much effort that the length of such a stretch cannot be much longer than M. Otherwise stated, this is the familiar statement that the longest q–atypical sub–stretch of $\omega_1, \dots, \omega_N$ is of typical length $\sim \log N/\Sigma(q)$. So $T(\omega)$ is for us the end–point of a q–atypical stretch of length approximately $(\log T(\omega))/\Sigma(q)$: by looking for sufficiently long q–atypical stretches we have always the freedom to choose $T(\omega) \gg 1$, in such a way that also $\log T(\omega) \ll T(\omega)$ and this is helpful for the estimates. So let us bound $Z_{T(\omega),\omega}$ from below by considering only the trajectories of the walk that stay in the upper half–plane up to the beginning of the q–atypical stretch and that are negative in the stretch, coming back to zero at step $T(\omega)$ (see Fig. 3: the polymer is cut at the first dashed vertical line). The contribution of these trajectories is easily evaluated: it is approximately

$$\left(\frac{1}{T(\omega)^{3/2}}\right) \exp\left(-2\lambda(q+h)\frac{\log T(\omega)}{\Sigma(q)}\right). \tag{3.7}$$

For such an estimate we have used (1.12) and $\log T(\omega) \ll T(\omega)$ both in writing the probability that the first return to zero of the walk is at the beginning of the q–atypical

stretch and in neglecting the probability that the walk is negative inside the stretch. It is straightforward to see that if

$$\frac{4\lambda}{3}h < -\frac{4\lambda}{3}q - \Sigma(q), \tag{3.8}$$

and if $T(\omega)$ is large, then also the quantity in (3.7) is large. We can still optimize this procedure by choosing q (which must be sufficently negative, i.e. $q < -h$). By playing with (3.6) one sees that one can choose $q_0 \in \mathbb{R}$ such that for $q = q_0$ the right–hand side in (3.8) equals $\log \mathsf{M}(-4\lambda/3)$ and if $h < \log \mathsf{M}(-4\lambda/3)/(4\lambda/3) = \underline{h}(\lambda)$ then $q_0 < -h$. This argument therefore is saying that there exists $C > 1$ such that

$$Z_{T(\omega),\omega}(0) \geq C, \tag{3.9}$$

for every ω. It only remains to show that $\mathbb{E}[T] < \infty$: this fact, together with a detailed proof of the argument just presented can be found in Appendix B.

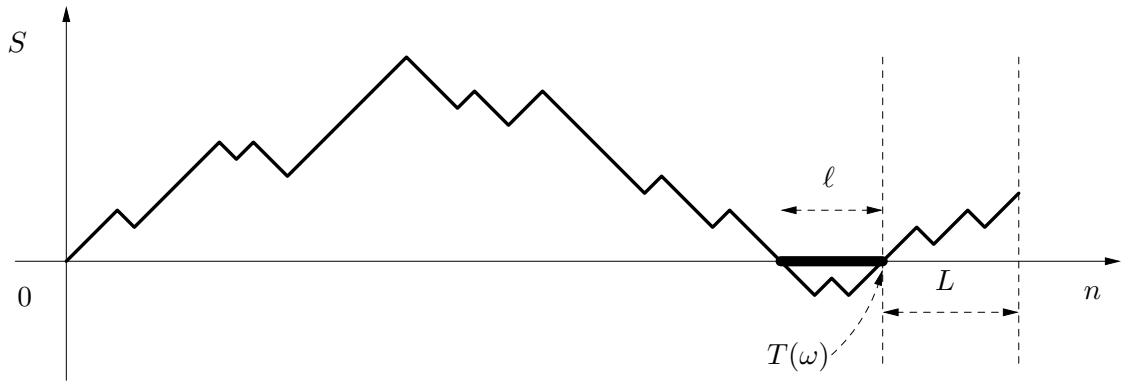

FIGURE 3. Inequality (3.10) comes simply from restricting the evaluation of $Z_{T(\omega)+L,\omega}$ to the trajectories visiting the q–*atypical* stretch of length ℓ and by staying away from the unfavorable solvent after that.

3.2. **Persistence of the effect of rare stretches.** As pointed out in the previous section, there is strong evidence that $h_c(\lambda) > \underline{h}(\lambda)$. At this stage Fig. 4 is of particular interest. Notice first of all that in spite of being substantially above $\underline{h}(\cdot)$ the copolymer appears to be still localized, see in particular case A.

The rigorous lower bounds that we are able to prove cannot establish localization in the region we are considering. All the same, notice that if one does not cut the polymer at $T(\omega)$, as in the argument above, but at $T(\omega) + L$, a lower bound of the following type

$$Z_{T(\omega)+L,\omega} \overset{\text{roughly}}{\geq} \text{const.}\frac{1}{T(\omega)^{3/2}} \exp\left(-2\lambda(q+h)\frac{\log T(\omega)}{\Sigma(q)}\right) \frac{1}{L^{1/2}}, \tag{3.10}$$

is easily established. Of course we are being imprecise, but we just want to convey the idea, see also Fig. 3, that after passing through an atypically *negative* stretch of environment ($q > 0$), the effect of this stretch decays at most like $L^{-1/2}$, that is the probability that a walk stays positive for a time L.

At this point we stress that the argument outlined in § 3.1 and re–used for (3.10) may be very well applied to $h > \underline{h}(\lambda)$, except that this time it does not suffice for (3.9). But it yields nevertheless that for $h \in \big(\underline{h}(\lambda), \overline{h}(\lambda)\big)$ the statement $Z_{N,\omega} \sim N^{-1/2}$, something a priori expected (for example [6]) in the delocalized regime and true for non disordered systems,

is violated. More precisely, one can find a sequence of random times $\{\tau_j\}_j$, $\lim_j \tau_j = \infty$ such that $Z_{\tau_j,\omega} \geq {\tau_j}^{-1/2+a}$, $a = a(\lambda, h) > 0$ (see Proposition 4.1 in [15]). These random times are constructed exactly by looking for q–atypical stretches as above and one can appreciate such an irregular decay for example in case B of Fig. 4, and this in spite of the fact that the data have been strongly coarse grained.

Therefore the lower bound (3.10), both in the localized and in the delocalized regime, yields the following picture: the lower bound we found on $Z_{N,\omega}$ grows suddenly in correspondence of atypical stretches and after that it decays with an exponent 1/2, up to another atypical stretch. This matches Fig. 4, at least at a qualitative level, see the caption of the figure.

4. The delocalized phase: a path analysis

Let us start with a qualitative observation: if we set the parameters (λ, h) of the copolymer to $(\lambda, h^{(m)}(\lambda))$ with $m = 0.9$, then the observed behavior of $\{Z^{\lambda,h}_{N,\omega}(0)\}_N$ –suitably averaged over blocks in order to eliminate local fluctuations– is somewhat close to $(\text{const})/N^{3/2}$. This is true for all the numerically accessible values of N (up to $N \sim 10^8$), at once for a number of values of λ and for a great number of typical environments ω. Of course this is suggesting that for $m = 0.9$ the curve $h^{(m)}(\lambda)$ lies in the delocalized region, but it is not easy to convert this qualitative observation into a precise statement, because we do not have a rigorous finite–volume criterion to state that a point (λ, h) belongs to the delocalized phase (the contrast with the localized phase, see (2.4), is evident). In other words, we cannot exclude the possibility that the system is still localized but with a characteristic size much larger than the one we are observing.

Nevertheless, the aim of this section is to give an empirical criterion, based on an analysis of the path behavior of the copolymer, that will allow us to provide some more quantitative argument in favor of the fact that the curve $h^{(m)}(\lambda)$ lies in the delocalized region even for values of $m < 1$. This of course would entail that the upper bound $\overline{h}(\lambda)$ defined in (1.7) is not strict.

4.1. Known and expected path behavior.

We want to look at the whole *profile* $\{Z^{\lambda,h}_{N,\omega^r}(x)\}_{x\in\mathbb{Z}}$ rather than only at $Z^{\lambda,h}_{N,\omega^r}(0)$, where by ω^r we mean the environment ω in the *backward direction*, that is $(\omega^r)_n := \omega_{N+1-n}$ (the reason for this choice is explained in Remark 4.1 below). The link with the path behavior of the copolymer, namely the law of S_N under the polymer measure $\mathbf{P}^{\lambda,h}_{N,\omega^r}$, is given by

$$\frac{Z^{\lambda,h}_{N,\omega^r}(x)}{Z^{\lambda,h}_{N,\omega^r}} = \mathbf{P}^{\lambda,h}_{N,\omega^r}(S_N = x)\,. \tag{4.1}$$

As already remarked in the introduction, although the localized and delocalized phases have been defined in terms of free energy they do correspond to sharply different path behaviors. In the localized phase it is known [24, 2] that the laws of S_N under $\mathbf{P}^{\lambda,h}_{N,\omega^r}$ are *tight*, which means that the polymer is essentially at $O(1)$ distance from the x–axis. The situation is completely different in the (interior of the) delocalized phase, where one expects that $S_N = O(\sqrt{N})$: in fact the conjectured path behavior (motivated by the analogy with the known results for non disordered models, see in particular [21], [11] and [8]) should be weak convergence under diffusive scaling to the *Brownian meander process* (that is Brownian motion conditioned to stay positive on the interval $[0, 1]$, see [23]). Therefore in

the (interior of the) delocalized phase the law of $S_N/\sqrt{N}$ under $\mathbf{P}^{\lambda,h}_{N,\omega^r}$ should converge weakly to the corresponding marginal of the Brownian meander, whose law has density $x\exp(-x^2/2)\mathbf{1}_{(x\geq 0)}$.

We stress however that for the delocalized regime the rigorous results that are available are more meager: essentially the only known $\mathbb{P}(\,\mathrm{d}\omega)$–a.s. result is that for any $L>0$

$$\lim_{N\to\infty}\mathbf{E}^{\lambda,h}_{N,\omega^r}\left[\frac{1}{N}\sum_{n=1}^{N}\mathbf{1}_{(S_n\geq L)}\right]=1\qquad \mathbb{P}(\,\mathrm{d}\omega)\text{–a.s.}\,, \tag{4.2}$$

that is the polymer spends almost all the time above any prefixed level. More precise results have been derived for the path behavior of the polymer under the *quenched averaged measure* $\mathbb{E}\mathbf{E}^{\lambda,h}_{N,\omega}[\,\cdot\,]$: these results go in the direction of proving the conjectured scaling limit, but they still do not suffice (we refer to [15] for more details and also for a discussion on what is still missing).

In spite of the lack of precise rigorous results, the analysis we are going to describe is carried out under the hypothesis that, in the interior of the delocalized phase, the scaling limit towards Brownian meander holds true (as it will be seen, the numerical results provide a sort of *a posteriori* confirmation of this hypothesis).

Remark 4.1. From a certain point of view attaching the environment backwards does not change too much the model: for example it is easy to check that if one replaces ω by ω^r in (1.3), the limit still exists $\mathbb{P}(\,\mathrm{d}\omega)$–a.s. and in $\mathbb{L}_1(\,\mathrm{d}\mathbb{P})$. Therefore the free energy is the same, because $\{\omega^r_n\}_{1\leq n\leq N}$ has the same law as $\{\omega_n\}_{1\leq n\leq N}$, for any fixed N.

However, if one focuses on the law of S_N as a function of N *for a fixed environment* ω, the behavior reveals to be much smoother under $\mathbf{P}^{\lambda,h}_{N,\omega^r}$ than under $\mathbf{P}^{\lambda,h}_{N,\omega}$. For instance, under the original polymer measure $\mathbf{P}^{\lambda,h}_{N,\omega}$ it is no more true that in the localized region the laws of S_N are tight (it is true only *most of the time*, see [14] for details). The reason for this fact is to be sought in the presence of long *atypical* stretches in every typical ω (this fact has been somewhat quantified in [15, Section 4] and it is at the heart of the approach in Section 3) that are encountered along the copolymer as N becomes larger. Of course the effect of these stretches is very much damped with the backward environment.

A similar and opposite phenomenon takes place also in the delocalized phase. In fancier words, we could say that for fixed ω and as N increases, the way S_N approaches its *limiting behavior* is faster when the environment is attached backwards: it is for this reason that we have chosen to work with $\mathbf{P}^{\lambda,h}_{N,\omega^r}$.

4.2. **Observed path behavior: a numerical analysis.** In view of the above considerations, we choose as a measure of the delocalization of the polymer the ℓ_1 distance $\triangle^{\lambda,h}_N(\omega)$ between the numerically computed profile for a polymer of size $2N$ under $\mathbf{P}^{\lambda,h}_{2N,\omega^r}$, and the conjectured asymptotic delocalized profile:

$$\triangle^{\lambda,h}_N(\omega):=\sum_{x\in 2\mathbb{Z}}\left|\frac{Z^{\lambda,h}_{N,\omega^r}(x)}{Z^{\lambda,h}_{N,\omega^r}}-\frac{1}{\sqrt{2N}}\,\varphi^+\!\left(\frac{x}{\sqrt{2N}}\right)\right|,\qquad \varphi^+(x):=x\,e^{-x^2/2}\mathbf{1}_{(x\geq 0)}\,. \tag{4.3}$$

Loosely speaking, when the parameters (λ,h) are in the interior of the the delocalized region we expect $\triangle_N$ to decrease to 0 as N increases, while this certainly will not happen if we are in the localized phase.

The analysis has been carried out at $\lambda=0.6$: we recall that the lower and upper bound of (1.7) give respectively $\underline{h}(0.6)\simeq 0.36$ and $\overline{h}(0.6)\simeq 0.49$, while the lower bound we

derived with our test for localization is $h = 0.41$, see Table 1. However, as observed in Section 3, Fig. 4, there is numerical evidence that $h = 0.43$ is still localized, and for this reason we have analyzed the values of $h = 0.44, 0.45, 0.46, 0.47$ (see below for an analysis on smaller values of h).

For each couple (λ, h) we have computed $\triangle_N^{\lambda,h}(\omega)$ for the sizes $N = a \times 10^6$ with $a = 1, 2, 5, 10$ and for 500 independent environments. Of course some type of statistical analysis must be performed on the data in order to decide whether there is a decay of $\triangle$ or not. The most direct strategy would be to look at the sample mean of a family of IID variables distributed like $\triangle_N(\omega)$, but it turns out that the fluctuations are too big to get reasonable confidence intervals for this quantity (in other words, the sample variance does not decrease fast enough), at least for the numerically accessible sample sizes. A more careful analysis shows that the variance is essentially due to a *very small* fraction of data that have *large* deviations from the mean, while the most of the data mass is quite concentrated.

For this reason we have chosen to focus on the *sample median* rather than on the sample mean. Table 3 contains the results of the analysis (see also Fig. 5 for a graphical representation): for each value of h we have reported the standard 95% confidence interval for the sample median (see Remark 4.2 below for details) for the four different values of N analyzed. While for $h = 0.44$ the situation is not clear, we see that for the values of h greater than 0.45 there are quantitative evidences for a decrease in $\triangle_N$: this leads us to the conjecture that the points (λ, h) with $\lambda = 0.6$ and $h \geq 0.45$ (equivalently, the points $(\lambda, h^{(m)}(\lambda))$ with $m \gtrsim 0.876$) lie in the delocalized region.

$h \backslash N (\times 10^6)$	1	2	5	10
0.44	[.0603, .0729]	[.0574, .0682]	[.0572, .0689]	[.0570, .0695]
0.45	[.0258, .0286]	[.0207, .0232]	[.0170, .0190]	[.0149, .0171]
0.46	[.0140, .0154]	[.0108, .0116]	[.00792, .00869]	[.00647, .00731]
0.47	[.00905, .00963]	[.00676, .00711]	[.00475, .00508]	[.00364, .00398]

TABLE 3. The table contains the standard 95% confidence interval for the median of a sample $\{\triangle_N^{\lambda,h}(\omega)\}_\omega$ of size 500, where $\lambda = 0.6$ and h, N take the different values reported in the table. For the values of $h \geq 0.45$ the decreasing behavior of $\triangle_N$ is quite evident (the confidence intervals do not overlap), see also Fig. 5.

As already remarked, these numerical observations cannot rule out the possibility that the system is indeed localized, but the system size is too small to see it. For instance, we have seen that there are evidences for $h = 0.43$ to be localized (see case C of Fig. 4). In any case, the exponential increasing of $Z_N(0)$ is detectable only at sizes of order$\sim 10^8$, while for smaller system sizes (up to$\sim 10^7$) the qualitative observed behavior of $Z_N(0)$ is rather closer to $(const)/N^{3/2}$, thus apparently suggesting delocalization (see case D of Fig. 4).

For this reason it is interesting to look at $\triangle_N^{0.6,h}$ for $h = 0.42, 0.43$ and for $N \ll 10^8$. For definiteness we have chosen $N = a \times 10^6$ with $a = 1, 2, 5, 10$, performing the computations for 3000 independent environments: the results are reported in Table 4 (see also Fig. 5). As one can see, this time there are clear evidences for an *increasing* behavior of $\triangle_N$. On the one hand this fact gives some more confidence on the data of Table 3, on the other hand it suggests that looking at $\{\triangle_N\}_N$ is a more reliable criterion for detecting (de)localization than looking at $\{Z_N(0)\}_N$.

$h\backslash N(\times 10^5)$	1	2	5	10
0.42	[.351, 0.382]	[.480, 0.517]	[.751, 0.794]	[1.01, 1.06]
0.43	[.143, 0.155]	[.165, 0.180]	[.197, 0.215]	[.236, 0.264]

TABLE 4. The table contains the standard 95% confidence interval for the median of a sample $\{\triangle_N^{\lambda,h}(\omega)\}_\omega$ of size 3000, where $\lambda = 0.6$ and h, N take the values reported in the table. For both values of h an increasing behavior of $\triangle_N$ clearly emerges, see also Fig. 5 for a graphical representation.

Remark 4.2. A confidence interval for the sample median can be obtained in the following general way (the steps below are performed under the assumption that the median is unique, which is, strictly speaking, not true in our case, but it will be clear that a finer analysis would not change the outcome). Let $\{Y_k\}_{1\le k\le n}$ denote a sample of size n, that is the variables $\{Y_k\}_k$ are independent with a common distribution, whose median we denote by $\xi_{1/2}$: $\mathbf{P}\left(Y_1 \le \xi_{1/2}\right) = 1/2$. Then the variable

$$\mathcal{N}_n := \#\{i \le n : \ Y_i \le \xi_{1/2}\} \tag{4.4}$$

has a binomial distribution $\mathcal{N}_n \sim B(n, 1/2)$ and when n is large (for us it will be at least 500) we can approximate $\mathcal{N}_n/n \approx 1/2 + Z/(2\sqrt{n})$, where $Z \sim N(0,1)$ is a standard gaussian. Let us denote the sample quantiles by Ξ_q, defined for $q \in (0,1)$ by

$$\#\{i \le n : \ Y_i \le \Xi_q\} = \lfloor qn \rfloor . \tag{4.5}$$

If we set $a := |\Phi^{-1}(0.025)|$ (Φ being the standard gaussian distribution function) then the random interval

$$\left[\Xi_{\frac{1}{2}-\frac{a}{2\sqrt{n}}}, \ \Xi_{\frac{1}{2}+\frac{a}{2\sqrt{n}}}\right] \tag{4.6}$$

is a 95% confidence interval for $\xi_{1/2}$, indeed

$$\begin{aligned} 0.95 = \mathbf{P}\big(Z \in [-a,a]\big) &= \mathbf{P}\left(\frac{1}{2} + \frac{1}{2\sqrt{n}} Z \ \in \ \left[\frac{1}{2} - \frac{a}{2\sqrt{n}} \ , \ \frac{1}{2} + \frac{a}{2\sqrt{n}}\right]\right) \\ \approx \mathbf{P}\left(\frac{\mathcal{N}_n}{n} \in \left[\frac{1}{2} - \frac{a}{2\sqrt{n}} \ , \ \frac{1}{2} + \frac{a}{2\sqrt{n}}\right]\right) &= \mathbf{P}\left(\Xi_{\frac{1}{2}-\frac{a}{2\sqrt{n}}} \le \xi_{1/2} \le \Xi_{\frac{1}{2}+\frac{a}{2\sqrt{n}}}\right). \end{aligned} \tag{4.7}$$

5. An empirical observation on the critical curve

The key point of this section is that, from a numerical viewpoint, $h_c(\cdot)$ seems very close to $h^{(m)}(\cdot)$, for a suitable value of m. Of course any kind of statement in this direction requires first of all a procedure to estimate $h_c(\cdot)$ and we explain this first.

Our analysis is based on the following conjecture:

$$(\lambda, h) \in \overset{\circ}{\mathcal{D}} \implies \lim_{N\to\infty} Z_{2N,\omega}^{\lambda,h}(0) \ = \ 0, \ \mathbb{P}\,(\,\mathrm{d}\omega) - \text{a.s.}. \tag{5.1}$$

The arguments in Section 3 (and in the Appendix) suggest the validity of such a conjecture, which is comforted by the numerical observation. Since, if $(\lambda, h) \in \mathcal{L}$, $Z_{2N,\omega}^{\lambda,h}(0)$ diverges (exponentially fast) $\mathbb{P}\,(\,\mathrm{d}\omega)$–almost surely and since $Z_{2N,\omega}^{\lambda,h}(0)$ is decreasing in h, we define $\hat{h}_{N,\omega}(\lambda)$ as the only h that solves $Z_{2N,\omega}^{\lambda,h}(0) = 1$. We expect that $\hat{h}_{N,\omega}(\lambda)$ converges to $h_c(\lambda)$ as N tends to infinity, for typical ω's. Of course setting the threshold to the value 1 is

rather arbitrary, but it is somewhat suggested by (2.4) and by the idea behind the proof of (3.1) (Proposition B.2 and equation (3.2)).

What we have observed numerically, see Figures 6 and 7, may be summed up by the statement

$$\text{there exists } m \text{ such that } \hat{h}_{N,\omega}(\lambda) \approx h^{(m)}(\lambda). \tag{5.2}$$

Practically this means that $\hat{h}_{N,\omega}(\lambda)$, for a set of λ ranging from 0.05 to 4, may be fitted with remarkable precision by the one parameter family of functions $\{h^{(m)}(\cdot)\}_m$. The fitting value of $m =: \hat{m}_{N,\omega}$ does depend on N and it is essentially increasing. This is of course expected since localization requires a sufficiently large system (recall in particular Table 2 and Fig. 2 – see the caption of Fig. 6 for the fitting criterion). We stress that we are presenting results that have been obtained for one fixed sequence of ω: based on what we have observed for example in Section 2.1 for different values of λ one does expect that for smaller values of λ one should use larger values of N, but changing N corresponds to selecting a longer, or shorter, stretch of ω, that is a different sequence of charges and this may have a rather strong effect on the value of $\hat{m}_{N,\omega}$. Moreover there is the problem of deciding which λ-dependence to choose. This may explain the deviations from (5.2) that are observed for small values of λ, but these are in any case rather moderate (see Fig. 7).

A source of stronger (and unavoidable) deviations arises in the cases of unbounded charges: of course if

$$h \geq h_{\text{sat}} := \max_{n\in\{1,\dots,N\}} \left(-(\omega_{2n-1}+\omega_{2n})/2\right), \tag{5.3}$$

then $Z^{\lambda,h}_{2N,\omega}(0) < 1$, regardless of the value of λ. Moreover it is immediate to verify that $\lim_{\lambda\to\infty} Z^{\lambda,h}_{2N,\omega}(0) = +\infty$ for $h < h_{\text{sat}}$ and therefore $\hat{h}_{N,\omega}(\lambda) \nearrow h_{\text{sat}}$ as $\lambda \nearrow \infty$. We refer to the captions of Fig. 7 for more on this saturation effect.

We have tried also alternative definitions of $\hat{h}_{N,\omega}(\lambda)$, namely:

(1) the value of h such that $Z^{\lambda,h}_{2N,\omega} = 1$ (or a different fixed value);
(2) the value of h such that the ℓ_1 distance between the distribution of the endpoint and the distribution of the meander, cf. Section 4, is smaller than a fixed threshold, for example 0.05.

What we have observed is that (5.2) still holds. What is not independent of the criterion is $\hat{m}_{N,\omega}$. Of course believing deeply in (5.2) entails the expectation that $\hat{m}_{N,\omega}$ converges to the non random quantity $h_c'(0)$. The results reported in this section suggest a value of $h_c'(0)$ larger than 0.83 and the cases presented in Section 4 suggest that it should be smaller than 0.86.

Appendix A. The algorithm for computing $Z_{N,\omega}$

We are going to briefly illustrate the algorithm we used in the numerical computation of the partition function $Z_N = Z^{\lambda,h}_{N,\omega}$. We recall its definition (see equation (1.1)):

$$Z_N = \mathbf{E}\left[\exp\left(-2\lambda\sum_{n=1}^{N}(\omega_n+h)\Delta_n\right)\right], \tag{A.1}$$

where $\Delta_n := (1-\text{sign}(S_n))/2$ and the convention for sign(0) described in the introduction.

Observe that a direct computation of Z_N from (A.1) would require to sum the contributions of 2^N random walk trajectories, making the problem numerically intractable. However, here we can make profitably use of the *additivity* of our Hamiltonian: loosely

speaking, if we join together two (finite) random walk segments, the energy of the resulting path is the sum of the energies of the building segments.

We can exploit this fact to derive a simple recurrence relation for the sequence of functions $\big\{\mathcal{Z}_M(y) := Z_{2M}(2y),\ y \in \mathbb{Z}\big\}_{M\in\mathbb{N}}$, where $Z_N(x) = Z^{\lambda,h}_{N,\omega}(x)$, the latter defined in (1.11), and we recall that we work with even values of N. Conditioning on S_{2M} and using the Markov property one easily finds

$$\mathcal{Z}_{M+1}(y) = \begin{cases} \frac{1}{4}\mathcal{Z}_M(y+1) \ + \ \frac{1}{2}\mathcal{Z}_M(y) \ + \ \frac{1}{4}\mathcal{Z}_M(y-1) & y>0 \\ \frac{1}{4}\Big[\mathcal{Z}_M(1)+\mathcal{Z}_M(0)\Big] \ + \ \frac{1}{4}\alpha_M\Big[\mathcal{Z}_M(0)+\mathcal{Z}_M(-1)\Big] & y=0 \\ \alpha_M\Big[\frac{1}{4}\mathcal{Z}_M(y+1) \ + \ \frac{1}{2}\mathcal{Z}_M(y) \ + \ \frac{1}{4}\mathcal{Z}_M(y-1)\Big] & y<0 \end{cases}, \tag{A.2}$$

where we have put $\alpha_M := \exp\big(-2\lambda\,(\omega_{2M+1}+\omega_{2M+2}+2h)\big)$.

From equation (A.2) and from the trivial observation that $\mathcal{Z}_M(y) = 0$ for $|y| > M$, it follows that $\{\mathcal{Z}_{M+1}(y),\ y \in \mathbb{Z}\}$ can be obtained from $\{\mathcal{Z}_M(y),\ y \in \mathbb{Z}\}$ with $O(M)$ computations. This means that we can compute Z_N in $O(N^2)$ steps.[1]

We point out that sometimes one is satisfied with *lower bounds* on Z_N, for instance in the statistical text for localization described in Section 2.1. In this case the algorithm can be further speeded up by restricting the computation to a suitable set of random walk trajectories. In fact when the system size is N the polymer is at most at distance $O(\sqrt{N})$ (we recall the discussion in Section 4 on the path behavior), hence a natural choice to get a lower bound on Z_N is to only take into account the contribution coming from those random walk paths $\{s_n\}_{n\in\mathbb{N}}$ for which

$$-A\sqrt{n} \le s_n \le B\sqrt{n} \qquad \text{for } n \ge N_0\,, \tag{A.3}$$

where A, B, N_0 are positive constants. Observe that this is easily implemented in the algorithm described above: it suffices to apply relation (A.2) only for $y \in [-A\sqrt{M}, B\sqrt{M}]$, while setting $\mathcal{Z}_{M+1}(y) = 0$ for the other values of y. In this way the number of computations needed to obtain Z_N is reduced to $O(N^{3/2})$.

The specific values of A, B, N_0 we used in our numerical computations are $3, 8, 1000$, and we would like to stress that the lower bound on Z_N we got coincides up to the 8th decimal digit with the *true value* obtained applying the complete algorithm.

A final important remark is that for the results we have reported we have used the Mersenne–Twister [19] pseudo–random number generator. However we have also tried other pseudo–random number generators and *true randomness* from `www.random.org`: the results appear not to depend on the generator.

Appendix B. Proof of the lower bound on h_c

We are going to give a detailed proof of the lower bound (3.1) on the critical curve, together with some related result. We stress that this appendix can be made substantially lighter if one is interested only in the *if* part of Proposition B.2. In this case the first part of this appendix is already contained in the first part of § 3.1, up to (3.4), and it suffices to look at § B.2.

[1]The algorithm just described can be implemented in a standard way: the code we used, written in C, is available on the web page: `http://www.proba.jussieu.fr/pageperso/giacomin/C/prog.html`. Graphic representations and standard statistical procedures have been performed with R [28].

We recall that $Z^{\lambda,h}_{N,\omega}(0)$ is the partition function corresponding to the polymer pinned at its right endpoint, see (1.11), and $T^C = T^C(\omega)$ is the first N for which $Z_{N,\omega}(0) \geq C$, see (3.2). In particular, for all ω such that $T^C(\omega) < \infty$ we have

$$Z^{\lambda,h}_{T^C(\omega),\omega}(0) \geq C\,. \tag{B.1}$$

We will also denote by $\mathcal{F}_n := \sigma(\omega_1, \dots, \omega_n)$ the natural filtration of the sequence $\{\omega_n\}_{n\in\mathbb{N}}$.

B.1. **A different look at (de)localization.** We want to show that (de)localization can be read from T^C. We introduce some notation: given an increasing, $2\mathbb{N}$–valued sequence $\{t_i\}_{i\in\mathbb{N}}$, we set $t_0 := 0$ and $\zeta_N := \max\{k : t_k \leq N\}$. Then we define

$$\begin{aligned}\widehat{Z}_{N,\omega}(0) = \widehat{Z}^{\{t_i\},\lambda,h}_{N,\omega}(0) \;&:=\; \mathbf{E}\left[e^{-2\lambda\sum_{n=1}^N(\omega_n+h)\Delta_n};\, S_{t_1}=0,\,\dots,\,S_{t_{\zeta_N}}=0,\,S_N=0\right]\\ &=\; \prod_{i=0}^{\zeta_N-1} Z^{\lambda,h}_{t_{i+1}-t_i,\theta^{t_i}\omega}(0)\;\cdot\; Z^{\lambda,h}_{N-t_{\zeta_N}(\omega),\theta^{t_{\zeta_N}}\omega}(0)\,,\end{aligned} \tag{B.2}$$

and we recall that θ denotes the translation on the environment. One sees immediately that $\widehat{Z}_{N,\omega}(0) \leq Z_{N,\omega}(0)$. We first establish a preliminary result.

Lemma B.1. *If the sequence $\{t_i\}_i$ is such that $\zeta_N/N \to 0$ as $N \to \infty$, then*

$$\lim_{N\to\infty}\frac{1}{N}\log \widehat{Z}^{\{t_i\},\lambda,h}_{N,\omega}(0) \;=\; \mathrm{F}(\lambda,h)\,, \tag{B.3}$$

both $\mathbb{P}(\,\mathrm{d}\omega)$–a.s. and in $\mathbb{L}_1(\mathbb{P})$.

Proof. By definition we have $Z_{N,\omega}(0) \geq \widehat{Z}_{N,\omega}(0)$. On the other hand, we are going to show that

$$Z^{\lambda,h}_{N,\omega}(0) \;\leq\; 4^{\zeta_N}\,A^{2\zeta_N}\left(\prod_{i=1}^{\zeta_N}(t_i-t_{i-1})\cdot(N-t_{\zeta_N})\right)^3\,\widehat{Z}^{\{t_i\},\lambda,h}_{N,\omega}(0)\,, \tag{B.4}$$

where A is a positive constant. To derive this bound, we resort to the equation (1.13) that expresses $Z_{N,\omega}(0)$ in terms of random walk excursions. We recall that $K(2n)$ is the discrete probability density of the first return time of the walk S to 0, and that $K(t) \geq 1/(A\,t^{3/2})$, $t \in 2\mathbb{N}$, for some positive constant A: it follows that for $a_1, \dots, a_k \in 2\mathbb{N}$

$$K(a_1+\ldots+a_k) \;\leq\; 1 \;\leq\; A^k\,(a_1\cdot\ldots\cdot a_k)^{3/2}\,K(a_1)\cdot\ldots\cdot K(a_k)\,. \tag{B.5}$$

This gives us an upper bound to the entropic cost needed to split a random walk excursion of length $(a_1+\ldots+a_k)$ into k excursions of lengths $a_1, \dots, a_k$.

Now let us come back to the second line of (1.13), that can be rewritten as

$$Z_{N,\omega}(0) \;=\; \sum_{\{x_i\}\subseteq\{0,\dots,N\}\cap 2\mathbb{N}} G(\{x_i\})\,. \tag{B.6}$$

A first observation is that if we restrict the above sum to the $\{x_i\}$ such that $\{x_i\} \supseteq \{t_i\}$, then we get $\widehat{Z}^{\{t_i\}}_{N,\omega}(0)$. Now for each $\{x_i\}$ we aim at finding an upper bound on the term $G(\{x_i\})$ of the form $c\cdot G(\{x_i\}\cup\{t_i\})$ for some $c > 0$ not depending on $\{x_i\}$. Each term $G(\{x_i\})$, see (1.13), is the product of two terms: an entropic part depending on $K(\cdot)$ and an energetic part depending on $\varphi(\cdot)$. Replacing the entropic part costs no more than

$$c_{\mathrm{ent}} \;:=\; A^{2\zeta_N}\left(\prod_{i=1}^{\zeta_N}(t_i-t_{i-1})\cdot(N-t_{\zeta_N})\right)^3\,, \tag{B.7}$$

thanks to (B.5). On the other hand, the cost for replacing the energetic part is easily bounded above by

$$c_{\text{energy}} := 2^{\zeta_N}\,, \tag{B.8}$$

so that the bound $G(\{x_i\}) \le c\cdot G(\{x_i\}\cup\{t_i\})$ holds true with $c := c_{\text{ent}}\, c_{\text{energy}}$. Replacing in this way each term in the sum in the r.h.s. of (B.6), we are left with a sum of terms $G(\{y_i\})$ corresponding to sets $\{y_i\}$ such that $\{y_i\} \supseteq \{t_i\}$. It remains to count the *multiplicity* of any such $\{y_i\}$, that is how many original sets $\{x_i\}$ are such that $\{x_i\} \cup \{t_i\} = \{y_i\}$. Sets $\{x_i\}$ satisfying this last condition must differ only for a subset of $\{t_i\}$, hence the sought multiplicity is 2^{ζ_N} (the cardinality of the parts of $\{t_i\}$) and the bound (B.4) follows.

Therefore we get

$$\begin{aligned}\left|\frac{\log \widehat{Z}^{\{t_i\},\lambda,h}_{N,\omega}(0)}{N} - \frac{\log Z^{\lambda,h}_{N,\omega}(0)}{N}\right| &\le (2\log 2A)\frac{\zeta_N}{N} \;+\; 3\,\frac{1}{N}\,\log\left(\prod_{i=1}^{\zeta_N}(t_i - t_{i-1})\cdot(N - t_{\zeta_N})\right) \\ &\le (2\log 2A)\frac{\zeta_N}{N} \;+\; 3\,\frac{\zeta_N+1}{N}\,\log\left(\frac{N}{\zeta_N+1}\right),\end{aligned} \tag{B.9}$$

where in the second inequality we have made use of the elementary fact that once the sum of k positive numbers is fixed, their product is maximal when all the numbers coincide (for us $k = \zeta_N + 1$). Since by hypothesis $\zeta_N/N \to 0$ as $N \to \infty$, the Lemma is proved. □

Now we are ready to prove the characterization of $\mathcal{L}$ and $\mathcal{D}$ in terms of T^C. Fix any $C > 1$.

Proposition B.2. *A point (λ, h) is localized, that is $h < h_c(\lambda)$, if and only if $\mathbb{E}[T^C] < \infty$.*

Proof. We set $\mathcal{A} := \{\omega : T^C(\omega) < \infty\}$. Observe that for $\omega \in \mathcal{A}^{\complement}$ we have $Z_{N,\omega}(0) \le C$ for every $N \in 2\mathbb{N}$, and consequently $\log Z^{\lambda,h}_{N,\omega}(0)/N \to 0$ as $N \to \infty$.

Consider first the case when the random variable T^C is defective, that is $\mathbb{P}[\mathcal{A}^{\complement}] > 0$ (this is a particular case of $\mathbb{E}[T^C] = \infty$). Since we know that $\log Z^{\lambda,h}_{N,\omega}(0)/N \to \mathrm{F}(\lambda, h)$, $\mathbb{P}(\,\mathrm{d}\omega)$–a.s., from the preceding observation it follows that $\mathrm{F}(\lambda, h) = 0$ and the Proposition is proved in this case.

Therefore in the following we can assume that T^C is proper, that is $\mathbb{P}(\mathcal{A}) = 1$, so that equation (B.1) holds for almost every ω. Setting $\theta^{-1}\mathcal{A} := \{\omega : \theta\omega \in \mathcal{A}\}$, we have $\mathbb{P}\left(\theta^{-1}\mathcal{A}\right) = 1$ since $\mathbb{P}$ is θ–invariant, and consequently $\mathbb{P}\left(\cap_{k=0}^{\infty}\theta^{-k}\mathcal{A}\right) = 1$, which amounts to saying that (B.1) can be actually strengthened to

$$Z^{\lambda,h}_{T^C(\theta^k\omega),\theta^k\omega}(0) \ge C \qquad \forall k \ge 0,\ \mathbb{P}(\,\mathrm{d}\omega)\text{–a.s.}\,. \tag{B.10}$$

Observe that the sequence $\{(\theta^{T^C(\omega)}\omega)_n\}_{n\in\mathbb{N}}$ has the same law as $\{\omega_n\}_{n\in\mathbb{N}}$ and it is independent of $\mathcal{F}_{T^C}$. We can define inductively an increasing sequence of stopping times $\{T_n\}_{n\in\mathbb{N}}$ by setting $T_0 := 0$ and $T_{k+1}(\omega) - T_k(\omega) := T^C(\theta^{T_k(\omega)}\omega) =: S_k(\omega)$. We also set $\zeta_N(\omega) := \max\{n :\ T_n(\omega) \le N\}$. Since $\{S_k\}_{k\in\mathbb{N}}$ is an IID sequence, by the strong law of large numbers we have that, $\mathbb{P}(\,\mathrm{d}\omega)$–a.s., $T_n(\omega)/n \to \mathbb{E}[T^C]$ as $n \to \infty$, and consequently $\zeta_N(\omega)/N \to 1/\mathbb{E}[T^C]$ as $N \to \infty$ (with the convention that $1/\infty = 0$).

Now let us consider the lower bound $\widehat{Z}_{N,\omega}(0)$ corresponding to the sequence $\{t_i\} = \{T_i(\omega)\}$: from (B.2) and (B.10) we get that $\mathbb{P}(\,\mathrm{d}\omega)$–a.s.

$$\begin{aligned}\widehat{Z}^{\{T_i(\omega)\},\lambda,h}_{N,\omega}(0) \;&=\; \prod_{i=0}^{\zeta_N(\omega)-1} Z^{\lambda,h}_{T^C(\theta^{T_i}\omega),\theta^{T_i}\omega}(0)\,\cdot\, Z^{\lambda,h}_{N-T_{\zeta_N(\omega)}(\omega),\theta^{T_{\zeta_N(\omega)}}\omega}(0)\\ &\geq\; C^{\zeta_N(\omega)}\cdot\frac{c}{N^{3/2}}\,,\end{aligned} \tag{B.11}$$

where c is a positive constant (to estimate the last term we have used the lower bound $Z_k(0) \geq c/k^{3/2}$, cf. (1.5)), and consequently

$$\mathrm{F}(\lambda,h) \;=\; \lim_{N\to\infty}\frac{\log Z^{\lambda,h}_{N,\omega}(0)}{N} \;\geq\; \liminf_{N\to\infty}\frac{\log \widehat{Z}^{\{T_i(\omega)\},\lambda,h}_{N,\omega}(0)}{N} \;\geq\; \frac{\log C}{\mathbb{E}[T^C]}\,. \tag{B.12}$$

It follows that if $\mathbb{E}[T^C] < \infty$ then $\mathrm{F}(\lambda,h) > 0$, that is (λ,h) is localized.

It remains to consider the case $\mathbb{E}[T^C] = \infty$, and we want to show that this time $\widehat{Z}_{N,\omega}(0)$, defined in (B.11), gives a null free energy. In fact, as $T^C(\eta)$ is defined as the *first* N such that $Z_{N,\eta}(0) \geq C$, it follows that $Z_{T^C(\eta),\eta}(0)$ cannot be much greater than C. More precisely, one has that

$$Z_{T^C(\eta),\eta}(0) \leq C\,\exp(2\lambda|\eta_{T^C(\eta)-1}+\eta_{T^C(\eta)}|)\,, \tag{B.13}$$

and from the first line of (B.11) it follows that

$$\frac{1}{N}\log\widehat{Z}_{N,\omega}(0) \;\leq\; \frac{\zeta_N(\omega)+1}{N}\log C \;+\; \frac{2\lambda}{N}\sum_{i=1}^{\zeta_N(\omega)}\Big(|\omega_{T_i(\omega)}|+|\omega_{T_i(\omega)-1}|\Big)\,. \tag{B.14}$$

We estimate the second term in the r.h.s. in the following way:

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{\zeta_N(\omega)}\Big(|\omega_{T_i(\omega)}|+|\omega_{T_i(\omega)-1}|\Big) &= \frac{1}{N}\sum_{k=1}^{N}\mathbf{1}_{\{\exists i:\,T_i(\omega)=k\}}\Big(|\omega_k|+|\omega_{k-1}|\Big)\\ &\leq \left(\frac{1}{N}\sum_{k=1}^{N}\mathbf{1}_{\{\exists i:\,T_i(\omega)=k\}}\right)^{1/2}\left(\frac{1}{N}\sum_{k=1}^{N}\Big(|\omega_k|+|\omega_{k-1}|\Big)^2\right)^{1/2}\\ &\leq \sqrt{\frac{\zeta_N(\omega)}{N}}\cdot 2\sqrt{\frac{1}{N}\sum_{k=1}^{N}|\omega_k|^2} \leq A\sqrt{\frac{\zeta_N(\omega)}{N}}\,,\end{aligned} \tag{B.15}$$

for some positive constant $A = A(\omega)$ and eventually as $N \to \infty$, having used the Cauchy–Schwartz inequality and the law of large numbers for the sequence $\{|\omega_k|^2\}_{k\in\mathbb{N}}$. Therefore

$$\frac{1}{N}\log\widehat{Z}_{N,\omega}(0) \;\leq\; \frac{\zeta_N(\omega)+1}{N}\log C \;+\; 4\lambda A\sqrt{\frac{\zeta_N(\omega)}{N}}\,, \tag{B.16}$$

and since $\mathbb{E}[T^C] = \infty$ implies $\zeta_N(\omega)/N \to 0$, $\mathbb{P}(\,\mathrm{d}\omega)$–a.s., we have $\log\widehat{Z}_{N,\omega}(0)/N \to 0$, $\mathbb{P}(\,\mathrm{d}\omega)$–a.s.. Then Lemma B.1 allows us to conclude that $\mathrm{F}(\lambda,h) = 0$, and the proof of the Proposition is completed. □

B.2. **Proof of the lower bound on h_c.** To prove equation (3.1), we are going to build, for every (λ, h) such that $h < \underline{h}(\lambda)$, a random time T such that $\mathbb{E}[T] < \infty$ and $Z^{\lambda,h}_{T(\omega),\omega}(0) \geq C$, for some $C > 1$. It follows that $T^C \leq T$, yielding that $\mathbb{E}[T^C] < \infty$ and by Proposition B.2 (λ, h) is localized, that is, $\underline{h}(\lambda) \leq h_c(\lambda)$.

Given $M \in 2\mathbb{N}$ and $q < -h$, we start defining the stopping time

$$\tau_M(\omega) = \tau_{M,q}(\omega) := \inf\left\{ n \in 2\mathbb{N} : \ \exists k \in 2\mathbb{N},\ k \geq M : \ \frac{\sum_{i=n-k+1}^{n} \omega_i}{k} \leq q \right\}. \tag{B.17}$$

This is the first instant at which a q–atypical stretch of length at least M appears along the sequence ω. The asymptotic behavior of τ_M is given by Theorem 3.2.1 in [10, § 3.2] which says that $\mathbb{P}(\,\mathrm{d}\omega)$–a.s.

$$\frac{\log \tau_M(\omega)}{M} \to \Sigma(q) \qquad \text{as } M \to \infty\,, \tag{B.18}$$

where $\Sigma(q)$ is Cramer's Large Deviations functional for ω, (3.6). We also give a name to the shortest of the terminal stretches in the definition of τ_M:

$$R_M(\omega) = R_{M,q}(\omega) := \inf\left\{ k \in 2\mathbb{N},\ k \geq M : \ \frac{\sum_{i=\tau_M-k+1}^{\tau_M} \omega_i}{k} \leq q \right\}, \tag{B.19}$$

and it is not difficult to realize that $R_M \leq 2M$.

We are ready to give a simple lower bound on the partition function of size $\tau_{M,q}$ (for any $M \in 2\mathbb{N}$ and $q < -h$): it suffices to consider the contribution of the trajectories that are negative in correspondence of the last (favorable) stretch of size R_M, and stay positive the rest of the time. Recalling that we use $K(\cdot)$ for the discrete density of the first return time to the origin and that by (1.12) we have $K(2n) \geq c/n^{-3/2}$ for a constant $c > 0$, we estimate

$$\begin{aligned} Z^{\lambda,h}_{\tau_M(\omega),\omega}(0) &\geq \frac{1}{4}\, K\left(\tau_M - R_M\right)\, K\left(R_M\right)\, e^{-2\lambda(q+h)R_M} \geq \frac{c^2}{4\tau_M^{3/2}(2M)^{3/2}} e^{-2\lambda(q+h)M} \\ &\geq c'\, \exp\left\{ \frac{3}{2} M \left[(-4\lambda/3)q - \frac{\log \tau_M}{M} - (4\lambda/3)h - \frac{\log M}{M} \right] \right\}, \end{aligned} \tag{B.20}$$

where $c' := c^2/(8\sqrt{2})$.

Having in mind (B.18), we define a random index $\ell = \ell_{A,\varepsilon,q}$ depending on the two parameters $A \in 2\mathbb{N},\ \varepsilon > 0$ and on q:

$$\ell(\omega) = \ell_{A,\varepsilon,q}(\omega) := \inf\left\{ k \in 2\mathbb{N},\ k \geq A : \ \frac{\log \tau_{k,q}(\omega)}{k} \leq \Sigma(q) + \varepsilon \right\}, \tag{B.21}$$

and we finally set $T(\omega) = T_{A,\varepsilon,q}(\omega) := \tau_{\ell(\omega)}(\omega)$. Then for the partition function of size $T(\omega)$ we get

$$Z^{\lambda,h}_{T(\omega),\omega}(0) \geq c'\, \exp\left\{ \frac{3}{2} A \left[(-4\lambda/3)q - \Sigma(q) - (4\lambda/3)h - \frac{\log A}{A} - \varepsilon \right] \right\}. \tag{B.22}$$

The fact that $\mathbb{E}[T_{A,\varepsilon,q}] < \infty$ for any choice of A, ε, q (with $q < -h$) is proved in Lemma B.3 below. It only remains to show that for every fixed (λ, h) such that $h < \underline{h}(\lambda)$, or equivalently

$$(4\lambda/3)h < \log \mathsf{M}(-4\lambda/3)\,, \tag{B.23}$$

the parameters A, ε, q can be chosen such that the right–hand side of equation (B.22) is greater than 1.

The key point is the choice of q. Note that the generating function $\mathsf{M}(\cdot)$ is smooth, since finite on the whole real line. Moreover for all $\lambda \in \mathbb{R}$ there exists some $q_0 \in \mathbb{R}$ such that

$$\log \mathsf{M}(-4\lambda/3) = (-4\lambda/3)q_0 - \Sigma(q_0)\,, \tag{B.24}$$

and from (B.23) it follows that $q_0 < -h$. Therefore we can take $q = q_0$, and equation (B.22) becomes

$$Z^{\lambda,h}_{T(\omega),\omega}(0) \geq c' \exp\left\{\frac{3}{2}A\left[\log \mathsf{M}(-4\lambda/3) - (4\lambda/3)h - \frac{\log A}{A} - \varepsilon\right]\right\}. \tag{B.25}$$

It is now clear that for every (λ, h), such that (B.23) holds, by choosing ε sufficiently small and A sufficiently large, the right–hand side of (B.25) is greater than 1, and the proof of (3.1) is complete.

Lemma B.3. *For every $A \in 2\mathbb{N}$, $\varepsilon > 0$ and $q < -h$ the random variable $T(\omega) = T_{A,\varepsilon,q}(\omega)$ defined below* (B.21) *is integrable: $\mathbb{E}[T] < \infty$.*

Proof. By the definition (B.21) of $\ell = \ell_{A,\varepsilon,q}$ we have

$$T_{A,\varepsilon,q} \leq \exp\left((\Sigma(q) + \varepsilon)\,\ell_{A,\varepsilon,q}\right), \tag{B.26}$$

so it suffices to show that for any $\beta > 0$ the random variable $\exp(\beta\,\ell_{A,\varepsilon,q})$ is integrable.

For any $l \in 2\mathbb{N}$, we introduce the IID sequence of random variables $\{Y^l_n\}_{n\in\mathbb{N}}$ defined by

$$Y^l_n := \frac{1}{l}\sum_{i=(n-1)l+1}^{nl} \omega_i\,. \tag{B.27}$$

By Cramer's Theorem [10] we have that for any fixed $q < 0$ and $\varepsilon > 0$ there exists l_0 such that $\mathbb{P}\left(Y^l_1 \leq q\right) \geq e^{-l(\Sigma(q)+\varepsilon/2)}$ for every $l \geq l_0$. By (B.21) have that

$$\{\ell > l\} \subseteq \{\tau_l > \exp((\Sigma(q) + \varepsilon)l)\} \subseteq \bigcap_{i=1}^{\lfloor M/l\rfloor} \{Y^l_i > q\}\,, \tag{B.28}$$

with $M := \exp((\Sigma(q) + \varepsilon)l)$, so that

$$\begin{aligned}\mathbb{P}\left(\ell > l\right) \leq \left(1 - e^{-l(\Sigma(q)+\varepsilon/2)}\right)^{\lfloor M/l\rfloor} &\leq \exp\left(-\lfloor M/l\rfloor e^{-l(\Sigma(q)+\varepsilon/2)}\right)\\ &\leq \exp\left(-\exp\left(l\varepsilon/4\right)\right),\end{aligned} \tag{B.29}$$

where the last step holds if l is sufficiently large (we have also used $1-x \leq e^{-x}$). Therefore

$$\mathbb{P}\left(\exp(\beta\,\ell) > N\right) = \mathbb{P}\left(\ell > (\log N)/\beta\right) \leq \exp\left(-N^{\varepsilon/4\beta}\right), \tag{B.30}$$

when N is large, and the proof is complete.

□

Acknowledgements

We thank Thierry Bodineau, Erwin Bolthausen and Fabio Toninelli for very useful discussions. We are also very grateful to Jacques Portes for the constant hardware assistance during the development of this work.

References

[1] S. Albeverio and X. Y. Zhou, *Free energy and some sample path properties of a random walk with random potential*, J. Statist. Phys. **83** (1996), 573–622.

[2] M. Biskup and F. den Hollander, *A heteropolymer near a linear interface*, Ann. Appl. Probab. **9** (1999), 668–687.

[3] E. Bolthausen and F. den Hollander, *Localization transition for a polymer near an interface*, Ann. Probab. **25** (1997), 1334–1366.

[4] T. Bodineau and G. Giacomin, *On the localization transition of random copolymers near selective interfaces*, J. Statist. Phys. **117** (2004), 801-818.

[5] E. Bolthausen and G. Giacomin, *Periodic copolymers at selective interfaces: a large deviations approach*, Ann. Appl. Probab. **15** (2005), 963–983.

[6] R. Bundschuh and T. Hwa, *Statistical mechanics of secondary structures formed by random RNA sequences*, Phys. Rev. E **65** (2002), 031903 (22 pages).

[7] F. Caravenna and G. Giacomin, *On constrained annealed bounds for linear chain pinning models*, preprint (2005).
`http://www.proba.jussieu.fr/pageperso/giacomin/pub/CG.pdf`

[8] F. Caravenna, G. Giacomin ad L. Zambotti, *A renewal theory approach to periodic inhomogeneous polymer models*, in preparation.

[9] M. S. Causo and S. G. Whittington, *A Monte Carlo investigation of the localization transition in random copolymers at an interface*, J. Phys. A: Math. Gen. **36**(2003), L189–L195.

[10] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Second Edition, Springer–Verlag, New York, 1998.

[11] J.–D. Deuschel, G. Giacomin and L. Zambotti, *Scaling limits of equilibrium wetting models in (1+1)–dimension*, Probab. Theory Rel. Fields
Online first: `link.springer-ny.com/link/service/journals/00440/`

[12] W. Feller, *An introduction to probability theory and its applications*, Vol. I, Third edition, John Wiley & Sons, Inc., New York–London–Sydney, 1968.

[13] D. S. Fisher, *Critical behavior of random transverse-field Ising spin chains*, Phys. Rev. B **51** (1995), 6411-6461.

[14] G. Giacomin, *Localization phenomena in random polymer models*, preprint (2004), available on the web page of the author.

[15] G. Giacomin and F. L. Toninelli, *Estimates on path delocalization for copolymers at interfaces*, accepted for publication on Probab. Theory Rel. Fields.

[16] T. Garel, D. A. Huse, S. Leibler and H. Orland, *Localization transition of random chains at interfaces*, Europhys. Lett. **8** (1989), 9–13.

[17] P. Le Doussal, C. Monthus and D. S. Fisher, *Random walkers in one-dimensional random environments: exact renormalization group analysis*, Phys. Rev. E (3) **59** (1999), 4795–4840.

[18] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, Vol. **89**, American Mathematical Society (2001).

[19] M. Matsumoto and T. Nishimura, *Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator*, ACM Trans. on Mod. and Comp. Simul. **8** (1998), 3–30.

[20] C. Monthus, *On the localization of random heteropolymers at the interface between two selective solvents*, Eur. Phys. J. B **13** (2000), 111–130.

[21] C. Monthus, T. Garel and H. Orland, *Copolymer at a selective interface and two dimensional wetting: a grand canonical approach*, Eur. Phys. J. B **17** (2000), 121–130.

[22] T. Morita, *Statistical mechanics of quenched solid solutions with application to magnetically dilute alloys*, J. Math. Phys. **5** (1966), 1401–1405.

[23] D. Revuz and M. Yor (1994), *Continuous martingales and Brownian motion*, 3rd ed., Springer-Verlag, Berlin.

[24] Ya. G. Sinai, *A random walk with a random potential*, Theory Probab. Appl. **38** (1993), 382–385.

[25] C. E. Soteros and S. G. Whittington, *The statistical mechanics of random copolymers*, J. Phys. A: Math. Gen. **37**, R279-R325.

[26] S. Stepanow, J.-U. Sommer and I. Ya. Erukhimovich, *Localization transition of random copolymers at interfaces*, Phys. Rev. Lett. **81** (1998), 4412–4416.

[27] A. Trovato and A. Maritan, *A variational approach to the localization transition of heteropolymers at interfaces*, Europhys. Lett. **46**(1999), 301–306.

[28] R Development Core Team, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria (2004), ISBN 3-900051-07-0.
URL http://www.R-project.org

Università di Milano-Bicocca, Dipartimento di Matematica e Applicazioni, Edificio U5, via Cozzi 53, 20125 Milano, Italy
and
Laboratoire de Probabilités de P 6 & 7 (CNRS U.M.R. 7599) and Université Paris 7 – Denis Diderot, U.F.R. Mathematiques, Case 7012, 2 place Jussieu, 75251 Paris cedex 05, France
E-mail address: f.caravenna@sns.it

Laboratoire de Probabilités de P 6 & 7 (CNRS U.M.R. 7599) and Université Paris 7 – Denis Diderot, U.F.R. Mathematiques, Case 7012, 2 place Jussieu, 75251 Paris cedex 05, France
Home page: http://www.proba.jussieu.fr/pageperso/giacomin/GBpage.html
E-mail address: giacomin@math.jussieu.fr

Dipartimento di Matematica Applicata "U. Dini", Università di Pisa, via Bonanno Pisano 25b, 56126 Pisa, Italy
E-mail address: m.gubinelli@dma.unipi.it

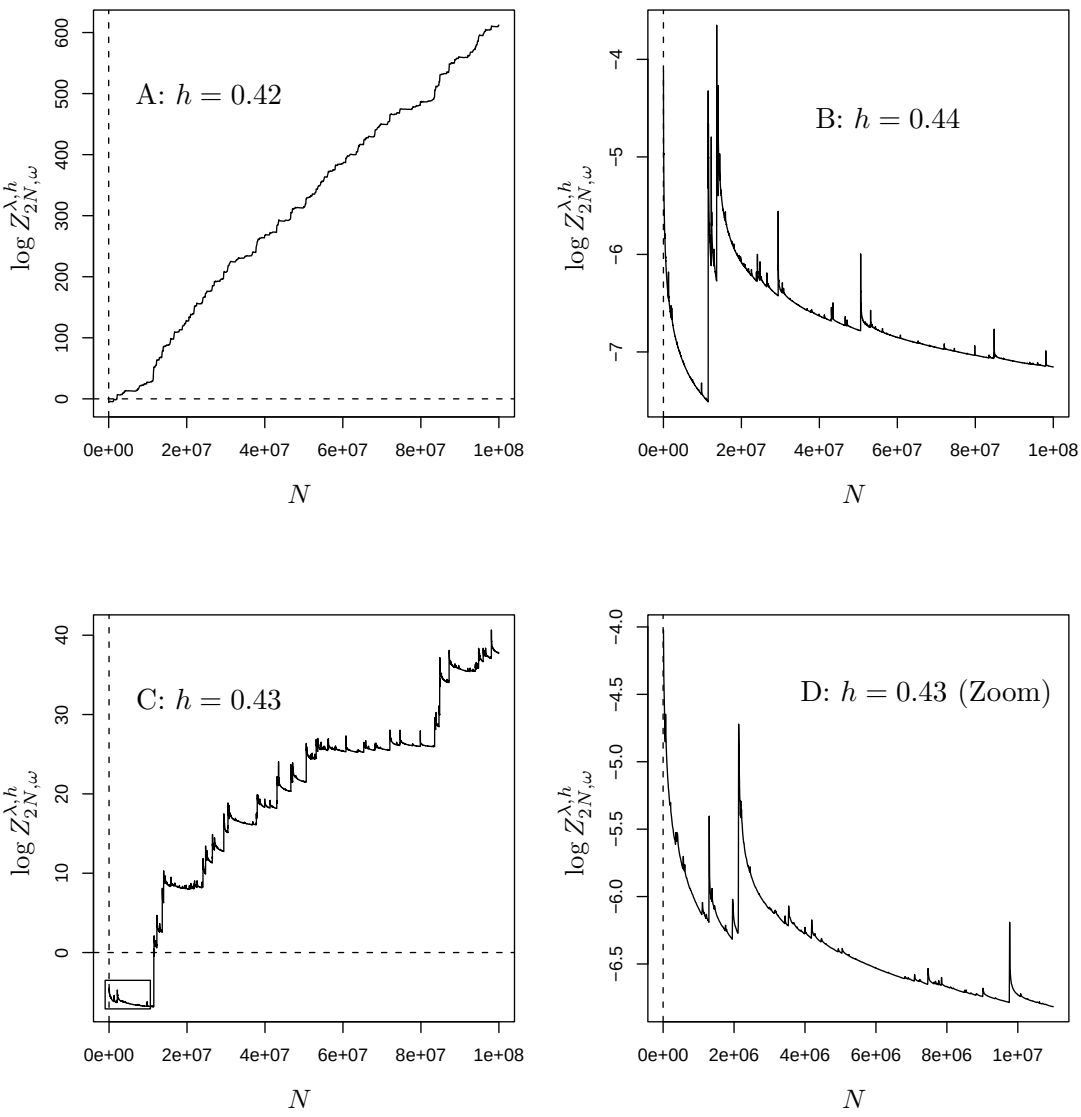

FIGURE 4. For $\lambda = 0.6$ ($\underline{h}(0.6) \simeq 0.36$ and $\overline{h}(0.6) \simeq 0.49$) , the behavior of $\log Z_{2N,\omega}$ for $h = 0.42$ (A), 0.43 (C,D) and 0.44 (B). The sequence of charges is the same in all the cases. In case A, the polymer is localized with free energy approximately $3 \cdot 10^{-6}$: the linear growth is quite clear, but a closer look shows sudden jumps, which correspond to atypically negative stretches of charges. Getting closer to the critical point, case C, the growth is still rather evident, but it is clearly the result of sudden growths followed by slow decays (approximately polynomial with exponent $-1/2$). Case B suggests delocalization: a closer analysis reveals a decay of the type $N^{-1/2}$, but sharp deviations are clearly visible and these deviations are in reality much larger, since in the graph we have plotted just one point every 10000. Case D is the zoom of the rectangle in the left corner of C. The similarity between B and D make clear that claiming delocalization looking at the behaviour of the partition function is difficult.

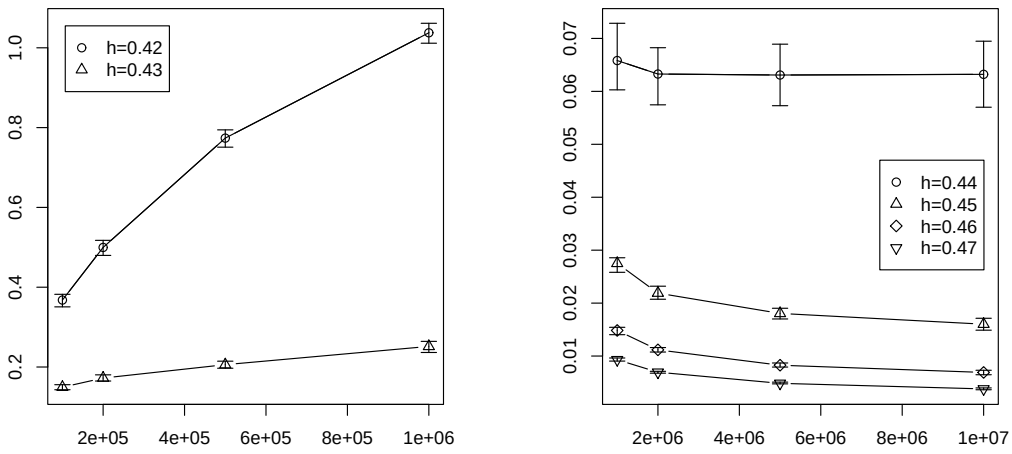

FIGURE 5. Graphical representation of the data of Tables 3 (on the right) and 4 (on the left). The plotted points are the sample medians against the sample size, the error bars correspond to the confidence intervals given in Tables 3 and 4.

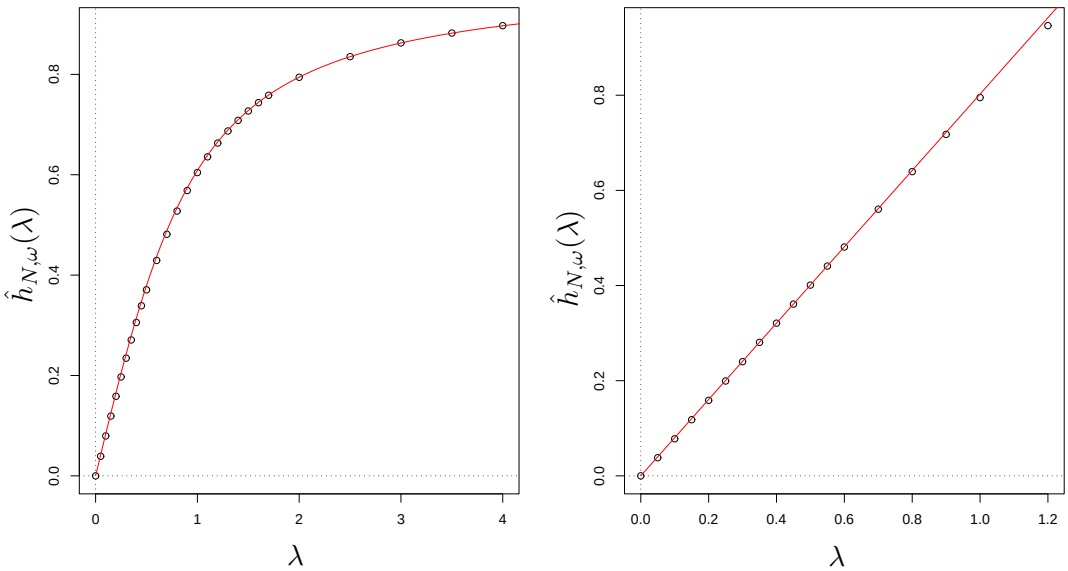

FIGURE 6. On the left the case of binary symmetric ω_1 and on the right the case of $\omega_1 \sim N(0,1)$, boths for $N = 3.2 \cdot 10^7$. The small circles represent the computed values: the errors on $\hat{h}_{N,\omega}(\lambda)$ are negligible and the plotted points are at the centers of the circles. The continuous line is instead the curve $h^{(m)}(\cdot)$. In the binary case $m = 0.841$ and it has been chosen by solving $h^{(m)}(4) = \hat{h}_{N,\omega}(4)$. In the Gaussian case $m = 0.802$, the maximum of $\hat{h}_{N,\omega}(\lambda)/\lambda$ for the plotted values of $\lambda(> 0)$. The rather different values of $\hat{m}_{N,\omega}$ may be somewhat understood both by considering that these two curves have been obtained for a fixed realization of ω and by taking into account the remark at the end of the caption of Table 2: it appears that for Gaussian charges one needs longer systems in order to get closer to the values of m observed in the binary case (in particular: for the prolongation, with the same random number generator, of the Gaussian ω sample used here up to $N = 5 \cdot 10^7$ one obtains $\hat{m}_{N,\omega} = 0.812$).

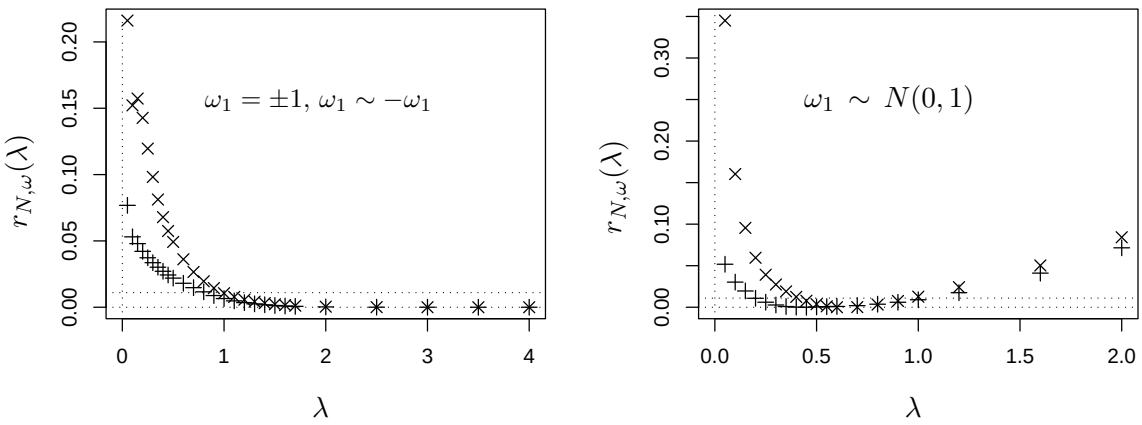

FIGURE 7. Relative errors $r_{N,\omega}(\lambda) := \left(h^{(m)}(\lambda) - \hat{h}_{N,\omega}(\lambda)\right)/\hat{h}_{N,\omega}(\lambda)$, for the value $m = \hat{m}_{N,\omega}$ explained in the caption of Fig. 6 and for the cases of $N = 2.5 \cdot 10^5$ ($\times$ dots), and $N = 3.2 \cdot 10^7$ ($+$ dots). Notice that in the binary case the error is more important for small values of λ (recall Table 2 and Fig. 2). Instead for the Gaussian case there is a deviation both for small and large values of λ: the deviation for large values is due to the saturation effect explained in the text. Given the fact that h_{sat}, cf. (5.3), behaves almost surely and to leading order for $N \to \infty$ as $\sqrt{\log N}$ one understand why the slow disappearing of the saturation effect has to be expected. In both graphs the dotted line above the axis is at level 0.01. The fitted values for $\hat{m}_{N,\omega}$, $N = 2.5 \cdot 10^5$, are 0.821 in the binary case and 0.778 in the Gaussian case.