RANDOM POLYMERS AND LOCALIZATION STRATEGIES

FRANCESCO CARAVENNA

Abstract. This is a **preliminary version** of the lecture notes for my course at the School on *Random Polymers and Related Topics*, Singapore, 14 - 25 May 2012. Warning: several parts are still to be completed and many references are missing. Feedback is very welcome.

Contents

1. Introduction

TO BE WRITTEN.

2. Pinning and copolymer models

2.1. **The building blocks.** The polymer models that we are going to study are built over two distinct sources of randomness, the *environment* (or *disorder*) $(\{\omega_n\}_{n\in\mathbb{N}}, \mathbb{P})$ and the *walk* $(\{S_n\}_{n\in\mathbb{N}}, \mathbf{P})$, that we will also call *polymer*. Let us state the precise assumptions we make on these random processes.

We assume that the environment $(\{\omega_n\}_{n\in\mathbb{N}}, \mathbb{P})$ is an i.i.d. sequence of real random variables, with zero mean and unit variance and with finite exponential moments:

$$\mathrm{M}(t) := \mathbb{E}(e^{t\omega_1}) < \infty\,, \qquad \forall t \in \mathbb{R}\,. \tag{2.1}$$

Note that $M(t) = 1 + \mathbb{E}(\omega_1)t + \frac{1}{2}\mathbb{E}(\omega_1^2) + o(t^2) = 1 + \frac{1}{2}t^2 + O(t^3)$ as $t \to 0$. Probably the most important example is the *Gaussian case*, when $\omega_1 \sim N(0,1)$, in which $M(t) = e^{t^2/2}$.

Date: May 13, 2012.

Remark 2.1. For most results it is enough that $\mathrm{M}(t) < \infty$ only in a neighborhood of $t = 0$, but for the sake of simplicity we always assume (2.1), except when otherwise mentioned. Another interesting generalization is to consider stationary and *ergodic* environment (cf. §7.3 below for the definition). We will mention (some of) the points where this is possible.

Next we come to the walk. Let $K : \mathbb{N} \to [0,1]$ and $\mathrm{T} \in \mathbb{N}$ be such that

$$\begin{aligned} &\sum_{n\in\mathbb{N}} K(n) \,=\, 1\,, \qquad\quad K(n) \,>\, 0 \quad \text{if and only if } n \in \mathrm{T}\mathbb{N}\,, \\ &K(n) \,=\, \frac{L(n)}{n^{1+\alpha}}\,,\ \forall n \in \mathrm{T}\mathbb{N}\,, \quad \text{with } \alpha \in [0,\infty),\ \ L : (0,\infty) \to (0,\infty) \ \textit{slowly varying}. \end{aligned} \tag{2.2}$$

Note that $K(\cdot)$ is (the discrete density of) a non defective probability on $\mathbb{N}$. We stress that the assumptions in (2.2) can be relaxed in several places, but we keep them for simplicity. Actually, in a few points we will even make stronger assumption, to lighten the exposition.

Remark 2.2. A function $L : (0,\infty) \to (0,\infty)$ is slowly varying if $\lim_{x\to\infty} L(tx)/L(x) = 1$ for every fixed $t > 0$. A complete reference on slowly varying function is [5]; for an essential compendium, more than sufficient for our purposes, we refer to [18, Appendix A.4]. The crucial fact to keep in mind is that a slowly varying function is asymptotically dominated by any polynomial:

$$\forall \varepsilon > 0\ \exists n_0 < \infty : \qquad n^{-\varepsilon} \,\le\, L(n) \,\le\, n^{\varepsilon}\,, \qquad \forall n \ge n_0\,,\ n \in \mathrm{T}\mathbb{N}\,. \tag{2.3}$$

A slightly more sophisticated feature is that as $N \to \infty$

$$\sum_{n=1}^{N} \frac{L(n)}{n^b} \sim \frac{L(N)}{(1-b)}\, N^{1-b} \quad \text{if } \ b < 1\,, \qquad \sum_{n=N}^{\infty} \frac{L(n)}{n^b} \sim \frac{L(N)}{(1-b)}\, \frac{1}{N^{b-1}} \quad \text{if } \ b > 1\,. \tag{2.4}$$

Remark 2.3. Typical examples of slowly varying functions are given by $L(x) \sim a(\log x)^b$ as $x \to \infty$, for any $a > 0$, $b \in \mathbb{R}$. The simplest and most important case is when the slowly varying function $L(\cdot)$ is *trivial*, meaning that

$$\lim_{x\to\infty} L(x) \,=\, c_K \in (0,\infty)\,. \tag{2.5}$$

We now state our hypotheses on the walk: $(\{S_n\}_{n\in\mathbb{N}}, \mathbf{P})$ is a Markov process on $\mathbb{Z}$ such that its zero level set is a renewal process with inter-arrival law $K(\cdot)$, and such that its excursions from zero lie entirely on one side of the axis, chosen by fair coin tossing:

$$\mathbf{P}(S_1 > 0, \ldots, S_{\mathrm{T}n-1} > 0, S_{\mathrm{T}n} = 0) = \mathbf{P}(S_1 < 0, \ldots, S_{\mathrm{T}n-1} < 0, S_{\mathrm{T}n} = 0) = \frac{K(\mathrm{T}n)}{2}\,, \tag{2.6}$$

for every $n \in \mathbb{N}$, except of course for the case $\mathrm{T} = 1$ and $n = 1$. In particular, by (2.4)

$$\mathbf{P}(S_1 > 0, \ldots, S_N > 0) \,=\, \sum_{n>N,\, n\in\mathrm{T}\mathbb{N}} \frac{K(n)}{2} \,\sim\, \frac{1}{\mathrm{T}}\, \frac{L(N)}{\alpha N^{\alpha}} \quad \text{as } N \to \infty\,. \tag{2.7}$$

For later convenience, we denote by $\{\tau_n\}_{n\in\mathbb{N}_0}$ the epochs of the successive visits of S to the state 0, that is

$$\tau_0 \,:=\, 0\,, \qquad \tau_{n+1} \,:=\, \inf\{k > \tau_n :\ S_k = 0\}\,, \tag{2.8}$$

so that $\{\tau_n\}_{n\in\mathbb{N}_0}$ is a renewal process with inter-arrival law $K(\cdot)$.

Remark 2.4. Basic examples of walks are given by symmetric nearest-neighbor random walks on $\mathbb{Z}$ with i.i.d. increments: if we set $p := 2\mathbf{P}(S_1 = +1) = 2\mathbf{P}(S_1 = -1) = 1 - \mathbf{P}(S_1 = 0) \in (0,1]$, then for $p = 1$ (the ordinary simple symmetric random walk) we have $\alpha = \frac{1}{2}$, $\mathrm{T} = 2$ and (2.5) holds with $c_K = \sqrt{2/\pi}$, cf. [17, §III.3], while for $0 < p < 1$ we have $\alpha = \frac{1}{2}$, $\mathrm{T} = 1$ and $c_K = \sqrt{p/(2\pi)}$, cf. [18, Proposition A.10].

Remark 2.5. Note that, given any probability $K(\cdot)$ satisfying (2.6), we can always build a (somewhat artificial and non nearest-neighbor) walk $(\{S_n\}_{n\in\mathbb{N}}, \mathbf{P})$ satisfying the above assumptions. In fact, if $\{\tau_n\}_{n\in\mathbb{N}_0}$ is a renewal process with inter-arrival law $K(\cdot)$ and $\{\sigma_n\}_{n\in\mathbb{N}}$ is an independent sequence of fair coin tosses, we can define

$$\gamma_n := \max\{k \in \mathbb{N}_0 : \tau_k \le n\}, \qquad S_n := \sigma_{\gamma_n} \cdot (n - \tau_{\gamma_n}),$$

that is S is the backward recurrence time process of τ multiplied by random signs.

Much more interesting examples are provided by Bessel-like random walks, cf. [1], that are nearest-neighbor Markov chains on $\mathbb{N}_0$ with local drift of the order $\approx x^{-1}$. Flipping their excursions through fair coin tossing, one obtains walks that satisfy (2.2) and (2.6) with $\mathrm{T} = 2$, for any $\alpha \ge 0$ and for a wide choice of $L(\cdot)$.

2.2. **The polymer partition function.** Let us fix an environment $(\omega = \{\omega_n\}_{n\in\mathbb{N}}, \mathbb{P})$ and a walk $(S = \{S_n\}_{n\in\mathbb{N}}, \mathbf{P})$ satisfying the assumptions of the previous section. Our main object of interest are the *pinning and copolymer partition functions* $Z^{\mathrm{pin}}_{N,\omega,\beta,h}$ and $Z^{\mathrm{cop}}_{N,\omega,\beta,h}$, defined for $N \in \mathbb{N}$, $\omega \in \mathbb{R}^{\mathbb{N}}$ and $\beta, h \in \mathbb{R}$ as the following positive functions:

$$Z^{\mathrm{pin}}_{N,\omega,\beta,h} = \mathbf{E}\Big[e^{\sum_{n=1}^N (\beta\omega_n - h)\mathbf{1}_{\{S_n=0\}}}\Big], \quad Z^{\mathrm{cop}}_{N,\omega,\beta,h} = \mathbf{E}\Big[e^{-2\beta\sum_{n=1}^N (\omega_n + h)\mathbf{1}_{\{S_{n-1}\le 0,\, S_n \le 0\}}}\Big]. \tag{2.9}$$

We usually restrict to $\beta \ge 0$, which entails no loss of generality (just consider $-\omega$). We are especially interested in the asymptotic behavior of these functions as $N \to \infty$, when ω is a $\mathbb{P}$-typical realization of the disorder. The interest of these objects will be explained in a moment.

To deal with both pinning and copolymer at the same time, we sometimes work with the following general partition function: for $N \in \mathbb{N}$, $\omega \in \mathbb{R}^{\mathbb{N}}$, and $\beta, h \in \mathbb{R}$

$$\mathsf{Z}_{N,\omega,\beta,h} = \mathsf{Z}^{\psi}_{N,\omega,\beta,h} = \mathbf{E}\Big[e^{\sum_{n=1}^N (\beta\omega_n - h)\psi(S_{[n-1,n]})}\Big], \tag{2.10}$$

so that for $\psi(x,y) = \psi^{\mathrm{pin}}(x,y) := \mathbf{1}_{\{y=0\}}$ we find precisely the pinning partition function. To recover the copolymer partition function, we have to take $\psi(x,y) = \psi^{\mathrm{cop}}(x,y) := \mathbf{1}_{\{x\le 0,\, y\le 0\}}$ and at the same time *we need to change the parametrization* ($\omega \to -\omega$, $\beta \to 2\beta$, $h \to 2\beta h$, or equivalently $\omega \to \omega$, $\beta \to -2\beta$, $h \to 2\beta h$):

$$\mathsf{Z}^{\psi^{\mathrm{pin}}}_{N,\omega,\beta,h} = Z^{\mathrm{pin}}_{N,\omega,\beta,h}, \qquad \mathsf{Z}^{\psi^{\mathrm{cop}}}_{N,\omega,-2\beta,2\beta h} = \mathsf{Z}^{\psi^{\mathrm{cop}}}_{N,-\omega,2\beta,2\beta h} = Z^{\mathrm{cop}}_{N,\omega,\beta,h}. \tag{2.11}$$

We should always keep in mind that, while the properties of the general partition function (2.10) are directly applicable to the pinning case, to recover the copolymer case it is necessary to perform the above change of variables.

Remark 2.6. One can study the partition function for general choices of $\psi : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$, beyond ψ^{pin} and ψ^{cop}, and indeed several interesting properties still hold (joint work in progress with Frank den Hollander). However, for the sake of simplicity, unless otherwise specified, we only focus on the cases $\psi = \psi^{\mathrm{pin}}$ and $\psi = \psi^{\mathrm{cop}}$.

The importance of the partition function is that it is a *generating function*. In fact it appears as the normalizing constant of the polymer measure $\mathsf{P}_{N,\omega,\beta,h}$, defined as:

$$\mathsf{P}_{N,\omega,\beta,h}(A) := \frac{1}{Z_{N,\omega,\beta,h}} \mathbf{E}\Big[e^{\sum_{n=1}^N (\beta\omega_n - h)\psi(S_{[n-1,n]})}\mathbf{1}_A\Big] =: \frac{\mathsf{Z}_{N,\omega,\beta,h}(A)}{\mathsf{Z}_{N,\omega,\beta,h}}. \tag{2.12}$$

The derivatives of the partition function $\mathsf{Z}_{N,\omega,\beta,h}$ are given by[†]

$$\begin{aligned}\frac{\partial}{\partial\beta}\log \mathsf{Z}_{N,\omega,\beta,h} &= \mathsf{E}_{N,\omega,\beta,h}\Bigg[\sum_{n=1}^{N}\omega_n\psi(S_{[n-1,n]})\Bigg],\\ \frac{\partial^2}{\partial\beta^2}\log \mathsf{Z}_{N,\omega,\beta,h} &= \mathsf{Var}_{N,\omega,\beta,h}\Bigg[\sum_{n=1}^{N}\omega_n\psi(S_{[n-1,n]})\Bigg] \geq 0.\end{aligned} \tag{2.13}$$

and analogously

$$\begin{aligned}\frac{\partial}{\partial h}\log \mathsf{Z}_{N,\omega,\beta,h} &= \mathsf{E}_{N,\omega,\beta,h}\Bigg[\sum_{n=1}^{N}-\psi(S_{[n-1,n]})\Bigg] \leq 0,\\ \frac{\partial^2}{\partial h^2}\log \mathsf{Z}_{N,\omega,\beta,h} &= \mathsf{Var}_{N,\omega,\beta,h}\Bigg[\sum_{n=1}^{N}\psi(S_{[n-1,n]})\Bigg] \geq 0.\end{aligned} \tag{2.14}$$

From these relations one can understand the central role of the partition function. For instance, understanding the properties of the derivatives of $\log \mathsf{Z}_{N,\omega,\beta,h}$ with respect to h gives information on the moments of the random variable $\sum_{n=1}^{N}\psi(S_{[n-1,n]})$ under the law $\mathsf{P}_{N,\omega,\beta,h}$. Therefore some essential properties of $\mathsf{P}_{N,\omega,\beta,h}$ as $N\to\infty$ are encoded in the asymptotic behavior of the partition function, to which we are going to devote our attention.

2.3. **The free energy.** The free energy $\mathsf{F}(\beta,h)=\mathsf{F}^{\psi}(\beta,h)$ is the rate of exponential growth of the partition function $\mathsf{Z}_{N,\omega,\beta,h}$, as the system size N diverges:

$$\mathsf{F}(\beta,h) := \lim_{N\to\infty}\frac{1}{N}\mathbb{E}\big[\log \mathsf{Z}_{N,\omega,\beta,h}\big] = \lim_{N\to\infty}\frac{1}{N}\log \mathsf{Z}_{N,\omega,\beta,h}, \quad \mathbb{P}\text{-a.s. and in } L^1(\mathbb{P}). \tag{2.15}$$

The existence of the limit (first equality) and the fact that it is self-averaging (second equality) are standard for the *constrained* partition function:

$$\mathsf{Z}^{\mathrm{c}}_{N,\omega,\beta,h} = \mathbf{E}\big[e^{\beta\sum_{n=1}^{N}(\omega_n-h)\psi(S_{[n-1,n]})}\,1_{\{S_N=0\}}\big], \qquad N\in \mathrm{T}\mathbb{N},$$

thanks to super-additivity, cf. [18, Ch. 4]. To transfer these results to the free partition function, one uses the following inequalities [18, eq. (4.25)]:

$$\mathsf{Z}^{\mathrm{c}}_{N,\omega,\beta,h} \leq \mathsf{Z}_{N,\omega,\beta,h} \leq (const.)\,N\,e^{-(\beta\omega_n-h)^-}\,\mathsf{Z}^{\mathrm{c}}_{N,\omega,\beta,h}, \qquad \forall N\in \mathrm{T}\mathbb{N},$$

and the intermediate values of $\mathrm{T}n < N < \mathrm{T}(n+1)$ cause no problem (exercise).

If we define the free energies $\mathrm{F}^{\mathrm{pin}}$ and $\mathrm{F}^{\mathrm{cop}}$ of the pinning and copolymer models in analogy with (2.15), but in terms of $Z^{\mathrm{pin}}_{N,\omega,\beta,h}$ and $Z^{\mathrm{cop}}_{N,\omega,\beta,h}$ respectively, by (2.11) we have

$$\mathrm{F}^{\mathrm{pin}}(\beta,h) = \mathsf{F}^{\psi^{\mathrm{pin}}}(\beta,h), \qquad \mathrm{F}^{\mathrm{cop}}(\beta,h) = \mathsf{F}^{\psi^{\mathrm{cop}}}(-2\beta,2\beta h), \tag{2.16}$$

(instead of -2β we could take 2β and at the same time $-\omega$ instead of ω).

Remark 2.7. We stress that the free energy $\mathsf{F}(\beta,h)$ does depend on *the law* $\mathbb{P}$ of the disorder (though it does not depend on the $\mathbb{P}$-typical realization of ω).

Remark 2.8. The existence and the self-averaging property of free energy, i.e. equation (2.15), still hold in the case of ergodic environment by Kingman's super-additive ergodic theorem, cf. [18, Ch. 4].

[†]The following computations are more conveniently done writing $Z_{\beta,h} := \mathbf{E}[e^{\beta X-hY}]$, for suitable X,Y.

The crucial observation is that

$$\mathsf{F}(\beta,h) \,\geq\, 0\,, \qquad \forall \beta \geq 0,\ h \in \mathbb{R}\,. \tag{2.17}$$

This is obtained from (2.15) by restricting the $\mathbf{E}$-expectation in $\mathsf{Z}_{N,\omega,\beta,h}$ to the event $\{S_1 > 0, \ldots, S_N > 0\}$, which by (2.7) has only polynomially vanishing probability. We then partition the (β, h) plane in the *localized* and *delocalized* regions:

$$\mathcal{L} \,:=\, \{(\beta,h) :\ \mathsf{F}(\beta,h) > 0\}\,, \qquad \mathcal{D} \,:=\, \{(\beta,h) :\ \mathsf{F}(\beta,h) = 0\}\,, \tag{2.18}$$

cf. Remark 2.10 below for some motivation. Characterizing the regions $\mathcal{L}$ and $\mathcal{D}$ is a very interesting problem, to which our efforts will be devoted.

2.4. **Convexity properties.** A crucial property of the free energy $\mathsf{F}(\beta, h)$ is that it is a *convex function* of (β, h):

$$\mathsf{F}\big(t(\beta_1,h_1) + (1-t)(\beta_2,h_2)\big) \,\leq\, t\,\mathsf{F}(\beta_1,h_1) \,+\, (1-t)\,\mathsf{F}(\beta_2,h_2)\,,$$

for all $(\beta_1, h_1), (\beta_2, h_2) \in \mathbb{R}^2$ and $t \in [0, 1]$. This is a simple consequence of Hölder's inequality (exercise). It follows in particular that $\mathsf{F}(\beta, h)$ is a continuous function, because it is finite for every (β, h): this follows implicitly from (2.15), but can also be verified directly (exercise).

In particular, $\mathsf{F}(\beta, h)$ is a convex function of each variable, when the other variable is kept fixed. By (2.16), the joint convexity in (β, h) is inherited by $\mathsf{F}^{\mathrm{pin}}$, but not by $\mathsf{F}^{\mathrm{cop}}$, due to the non-linear change of variables. However, the marginal convexity in each variable does hold for $\mathsf{F}^{\mathrm{cop}}$ as well, because the change of variables is linear in each variable.

Remark 2.9. Recalling (2.15), the convexity of $\mathsf{F}(\beta, h)$ in each variable can be also extracted from (2.13) and (2.14), which show that this property holds for $\frac{1}{N}\log \mathsf{Z}_{N,\omega,\beta,h}$ for every fixed N, noting that the limit of convex functions is convex.

We recall that a (finite) convex function $f : \mathbb{R} \to \mathbb{R}$ is continuous and is differentiable outside an exceptional set which is at most countable (and possibly empty, of course). More precisely, for every $t \in \mathbb{R}$ the left and right derivatives exist and are monotone functions:

$$\frac{\mathrm{d}f(s)}{\mathrm{d}s^-} \leq \frac{\mathrm{d}f(s)}{\mathrm{d}s^+} \leq \frac{\mathrm{d}f(t)}{\mathrm{d}t^-} \leq \frac{\mathrm{d}f(t)}{\mathrm{d}t^+}\,, \qquad \forall s \leq t\,.$$

Now let $f_n : \mathbb{R} \to \mathbb{R}$ be a sequence of convex functions that converge pointwise to the (necessarily convex) function $f : \mathbb{R} \to \mathbb{R}$. Then we almost have the convergence of the derivatives (exercise):

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t^-} \,\leq\, \liminf_{n\to\infty} \frac{\mathrm{d}f_n(t)}{\mathrm{d}t^-} \,\leq\, \limsup_{n\to\infty} \frac{\mathrm{d}f_n(t)}{\mathrm{d}t^+} \,\leq\, \frac{\mathrm{d}f(t)}{\mathrm{d}t^+}\,.$$

In particular, at every point t in which the derivatives of f_n and f exist (for every $n \in \mathbb{N}$), we have

$$\lim_{n\to\infty} \frac{\mathrm{d}f_n(t)}{\mathrm{d}t} \,=\, \frac{\mathrm{d}f(t)}{\mathrm{d}t}\,.$$

Recalling (2.14) and setting

$$\mathcal{L}_N := \sum_{n=1}^{N} \psi(S_{[n-1,n]})\,, \tag{2.19}$$

it follows from the above considerations that, for every (β, h) at which the partial derivative of $\mathsf{F}(\beta, h)$ with respect to h exists — that is, for any given β, for every h except for at most a countable set of points — we have for $\mathbb{P}$-a.e. ω

$$-\frac{\partial}{\partial h}\mathsf{F}(\beta, h) \;=\; \lim_{N\to\infty} \mathsf{E}_{N,\omega,\beta,h}\left[\frac{\mathcal{L}_N}{N}\right], \tag{2.20}$$

and in general

$$\frac{\partial}{\partial h^-}\mathsf{F}(\beta, h) \;\le\; \liminf_{N\to\infty} \mathsf{E}_{N,\omega,\beta,h}\left[-\frac{\mathcal{L}_N}{N}\right] \;\le\; \limsup_{N\to\infty} \mathsf{E}_{N,\omega,\beta,h}\left[-\frac{\mathcal{L}_N}{N}\right] \;\le\; \frac{\partial}{\partial h^+}\mathsf{F}(\beta, h)\,. \tag{2.21}$$

These relation show that the properties of (the derivatives of) the free energy $\mathsf{F}(\beta, h)$ give us important information on the properties of the *order parameter* $\frac{\mathcal{L}_N}{N}$ under the polymer measure $\mathsf{P}_{N,\omega,\beta,h}$. More details are provided in §2.7 below.

2.5. **Monotonicity properties.** By (2.14), for every $N \in \mathbb{N}$ the function $h \mapsto \frac{1}{N}\log \mathsf{Z}_{N,\omega,\beta,h}$ is non-increasing, hence by (2.15) $h \mapsto \mathsf{F}(\beta, h)$ is non increasing too, for every fixed $\beta \in \mathbb{R}$. For the monotonicity in β, note that by (2.13) for $\beta = 0$, averaging over ω we can write

$$\frac{\partial}{\partial \beta}\mathbb{E}\left[\frac{1}{N}\log \mathsf{Z}_{N,\omega,0,h}\right] \;=\; \mathbb{E}\mathsf{E}_{N,\omega,0,h}\left[\frac{1}{N}\sum_{n=1}^{N}\omega_n \psi(S_{[n-1,n]})\right] \;=\; 0\,, \qquad \forall N \in \mathbb{N}\,,$$

because for $\beta = 0$ the law $\mathsf{P}_{N,\omega,0,h} = \mathsf{P}_{N,0,0,h}$ does not depend on ω, hence the $\mathbb{E}$-expectation can be brought inside. Since $\beta \mapsto \frac{1}{N}\log \mathsf{Z}_{N,\omega,\beta,h}$ is a convex function with zero derivative at $\beta = 0$, it must be non-decreasing for $\beta \in [0, \infty)$. Therefore by (2.15) the function $\beta \to \mathsf{F}(\beta, h)$ is non-decreasing as well, for every fixed $h \in \mathbb{R}$.

Summarizing: the free energy $\mathsf{F}(\beta, h)$ is non-increasing in $h \in \mathbb{R}$ and non-decreasing in $\beta \in [0, \infty)$.

Remark 2.10. Recall (2.17) and the definition (2.18) of the regions $\mathcal{L}$ and $\mathcal{D}$.

If (β, h) is in $\mathcal{L}$, then by convexity and monotonicity $\frac{\partial}{\partial h^+}\mathsf{F}(\beta, h) < 0$ and by (2.21) the the order parameter $\frac{1}{N}\mathcal{L}_N$ averaged over the polymer measure $\mathsf{P}_{N,\omega,\beta,h}$ has a *strictly positive inferior limit*; furthermore, the inferior limit is a true limit whenever the partial derivative $\frac{\partial}{\partial h}\mathsf{F}(\beta, h)$ exists, by (2.20). Note that this is a clear localization path statement, because $\frac{1}{N}\mathcal{L}_N$ is the average number of monomers that are on the interface (in the pinning case) or below the interface (in the copolymer case).

On the other hand, if (β, h) is in the interior of $\mathcal{D}$, then $\frac{\partial}{\partial h}\mathsf{F}(\beta, h) = 0$, because F is constant in $\mathcal{D}$, hence by (2.19) and Markov's inequality the order parameter $\frac{1}{N}\mathcal{L}_N$ converges to zero in $\mathsf{P}_{N,\omega,\beta,h}$-probability. This is a delocalization path statement.

Summarizing, the definition of localization and delocalization given in terms of the free energy, which is quite indirect, has in reality a close link to path properties in terms of the order parameter $\frac{\mathcal{L}_N}{N}$. More details in the localized regions are given in §2.7.

2.6. **The homogeneous free energy.** The non-disodered case $\beta = 0$ (which is equivalent to $\omega = 0$) can be solved exactly. The crucial result is that $\mathsf{F}^{\psi}(0, h) = 0$ if $h \ge 0$ while $\mathsf{F}^{\psi}(0, h) > 0$ if $h < 0$, both for $\psi = \psi^{\mathrm{cop}}$ and $\psi = \psi^{\mathrm{pin}}$.

In the copolymer case, the free energy is actually totally explicit - and somewhat trivial:

$$\mathsf{F}^{\psi^{\mathrm{cop}}}(0, h) \;=\; 2\,|h|\,1_{\{h<0\}}\,, \qquad \forall h \in \mathbb{R}\,. \tag{2.22}$$

Note in particular that $\mathsf{F}^{\psi^{\mathrm{cop}}}(0,h)$ carries no dependence of the underlying walk. The proof is simple. For $h \le 0$, just observe that

$$e^{-hN}\,\mathbf{P}(S_1 < 0,\ldots,S_N < 0) \;\le\; \mathsf{Z}^{\psi^{\mathrm{cop}}}_{N,\omega,\beta,h} \;\le\; e^{-hN}\,,$$

and analogously, for $h \ge 0$,

$$\mathbf{P}(S_1 > 0,\ldots,S_N > 0) \;\le\; \mathsf{Z}^{\psi^{\mathrm{cop}}}_{N,\omega,\beta,h} \;\le\; 1\,.$$

Taking "$\frac{1}{N}\log$" in these relations and letting $N \to \infty$, recall that $\mathbf{P}(S_1 < 0,\ldots,S_N < 0) = \mathbf{P}(S_1 > 0,\ldots,S_N > 0)$ decays only polynomially in N, cf. (2.7), we find (2.22).

In the pinning case, the free energy $\mathsf{F}^{\psi^{\mathrm{cop}}}(0,h)$ is given by a less straightforward and more interesting formula. More precisely, $\mathsf{F}^{\psi^{\mathrm{cop}}}(0,h) = 0$ for $h \ge 0$, while for $h < 0$ it is given by the solution of the following implicit (but quite explicit) equation:

$$\mathsf{F}^{\mathrm{pin}}(0,h) = \lambda_h \quad \text{such that} \quad \sum_{n\in\mathrm{T}\mathbb{N}} K(n)\,e^{-\lambda_h n} \;=\; e^{-|h|}\,, \qquad \forall h < 0\,. \tag{2.23}$$

From this, using a Riemann sum approximation, the asymptotic behavior of for small h can be extracted. The net result is as follows:

$$\mathsf{F}^{\mathrm{pin}}(0,h) \;\sim\; \hat{L}\left(\frac{1}{|h|}\right)|h|^{1/\min\{\alpha,1\}}\,, \qquad \text{as } h \uparrow 0\,.$$

where $\hat{L}(\cdot)$ is a slowly varying function depending on $L(\cdot)$ and α, cf. [18, Theorem 2.1] for details (in the non-periodic case $\mathrm{T} = 1$).

Let us sketch the proof (2.23). Let us set for $h < 0$

$$\widetilde{K}_h(n) \;:=\; K(n)\,e^{-\lambda_h n}\,e^{-h}\,,$$

where λ_h is defined implicitly in (2.23), so that $\sum_{n\in\mathbb{N}} \widetilde{K}_h(n) = 1$. We are left with showing that $\lambda_h = \mathsf{F}^{\mathrm{pin}}(0,h)$. To this purpose, decompose the constrained partition function summing over the location of the return times to zero:

$$\begin{aligned}\mathsf{Z}^{\psi^{\mathrm{pin}},\mathrm{c}}_{N,\omega,\beta,h} &= \sum_{k=1}^{N}\;\sum_{0=:t_0<t_1<\ldots<t_k:=N}\;\prod_{i=1}^{k} e^{-h}\,K(t_i - t_{i-1}) \\ &= e^{\lambda_h N}\sum_{k=1}^{N}\;\sum_{0=:t_0<t_1<\ldots<t_k:=N}\;\prod_{i=1}^{k}\widetilde{K}_h(t_i - t_{i-1}) \;=\; e^{\lambda_h N}\,\widetilde{\mathbf{P}}_h(N\in\tau)\,,\end{aligned} \tag{2.24}$$

where we denote by $\widetilde{\mathbf{P}}_h$ the law under which τ is a renewal process with inter-arrival law $\widetilde{K}_h(\cdot)$. Note that $\sum_{n\in\mathbb{N}} k\widetilde{K}_h(n) < \infty$, hence the renewal τ has finite mean. By the classical renewal theorem, it follows that

$$\lim_{N\to\infty,\,N\in\mathrm{T}\mathbb{N}} \widetilde{\mathbf{P}}_h(N\in\tau) \;=\; \frac{1}{\sum_{n\in\mathbb{N}} n\,\widetilde{K}_h(n)} \in (0,\infty)\,.$$

Taking "$\frac{1}{N}\log$" in (2.24) and letting $N \to \infty$, we find $\lambda_h = \mathsf{F}^{\psi^{\mathrm{pin}}}(0,h)$.

Remark 2.11. In the pinning case, when $K(\infty) > 0$, **TO BE COMPLETED**.

2.7. **Path properties.** The characterization of localization and delocalization in terms of path properties, outlined in Remark 2.10, can be substantially strengthened, especially in the localized region.

An easy general improvement is as follows. Let us take a point $(\beta, h) \in \mathcal{L}$, that is $h < \mathsf{h}_c(\beta)$, such that the partial derivative $m_{\beta,h} := \frac{\partial}{\partial h}\mathsf{F}(\beta, h)$ exists. Then we can rewrite (2.20) as

$$\lim_{N\to\infty} \mathsf{E}_{N,\omega,\beta,h}\left[\frac{\mathcal{L}_N}{N}\right] \;=\; -m_{\beta,h}\,. \tag{2.25}$$

By a Taylor expansion

$$\mathsf{F}(\beta, h+\delta) \;=\; \mathsf{F}(\beta,h) \;+\; m_{\beta,h}\,\delta \;+\; o(\delta^2)\,, \qquad \text{as } \delta \to 0\,. \tag{2.26}$$

Fix $\varepsilon > 0$ and choose $\delta > 0$ small enough so that the $o(\delta^2)$ term above is smaller than $\frac{1}{3}\varepsilon\delta$. By Markov's inequality, we can write

$$\begin{aligned}\mathsf{P}_{N,\omega,\beta,h}\left(\frac{\mathcal{L}_N}{N} > -m_{\beta,h} + \varepsilon\right) \;&=\; \mathsf{P}_{N,\omega,\beta,h}\big(e^{\delta\mathcal{L}_N} > e^{\delta(-m_{\beta,h}+\varepsilon)N}\big) \;\le\; \frac{\mathsf{E}_{N,\omega,\beta,h}\big[e^{\delta\mathcal{L}_N}\big]}{e^{\delta(-m_{\beta,h}+\varepsilon)N}} \\ &=\; \exp\left(N\left(\frac{1}{N}\log \mathsf{Z}_{N,\omega,\beta,h-\delta} - \frac{1}{N}\log \mathsf{Z}_{N,\omega,\beta,h} + m_{\beta,h}\delta - \delta\varepsilon\right)\right).\end{aligned}$$

By (2.15), for every $\eta > 0$ and for $\mathbb{P}$-a.e. ω, there exists $N_0 = N_0(\omega)$ such that for all $N \ge N_0$

$$\frac{1}{N}\log \mathsf{Z}_{N,\omega,\beta,h-\delta} - \frac{1}{N}\log \mathsf{Z}_{N,\omega,\beta,h} + m_{\beta,h}\delta \;\le\; \mathsf{F}(\beta, h-\delta) - \mathsf{F}(\beta,h) + m_{\beta,h}\delta + \eta\,.$$

Choosing $\eta = \frac{1}{3}\varepsilon\delta$ and recalling (2.26) and the following line, it follows that for all $N \ge N_0$

$$\mathsf{P}_{N,\omega,\beta,h}\left(\frac{\mathcal{L}_N}{N} > -m_{\beta,h} + \varepsilon\right) \;\le\; \exp\left(-\frac{1}{3}\delta\varepsilon N\right).$$

With symmetrical arguments, we finally arrive at

$$\mathsf{P}_{N,\omega,\beta,h}\left(\left|\frac{\mathcal{L}_N}{N} - (-m_{\beta,h})\right| > \varepsilon\right) \;\le\; 2\,\exp\left(-\frac{1}{3}\delta\varepsilon N\right),$$

which shows that the order parameter $\frac{\mathcal{L}_N}{N}$ has strong concentration properties around its asymptotic limit $-m_{\beta,h}$ (recall (2.20)), with exponentially decaying probability.

TO BE COMPLETED

3. The phase diagram

3.1. **The critical curve.** Recall (2.17) and the definition (2.18) of the regions $\mathcal{L}$ and $\mathcal{D}$. Since $h \mapsto \mathsf{F}(\beta, h)$ is non-increasing, the regions $\mathcal{L}$ and $\mathcal{D}$ are separated by a *critical curve*

$$\mathsf{h}_c(\beta) \;=\; \mathsf{h}_c^{\psi}(\beta) \;:=\; \inf\{h \in \mathbb{R} :\; \mathsf{F}(\beta,h) = 0\} \;=\; \sup\{h \in \mathbb{R} :\; \mathsf{F}(\beta,h) > 0\}\,. \tag{3.1}$$

By the analysis of the homogeneous model in §2.6 we know that $\mathsf{F}(0,h) > 0$ for $h < 0$ and $\mathsf{F}(0,h) = 0$ for $h \ge 0$, hence $\mathsf{h}_c(0) = 0$.

Note that we can write $\mathcal{D} = \{(\beta,h) :\; \mathsf{F}(\beta,h) \le 0\}$ as a lower level set of $\textsc{f}$, hence $\mathcal{D}$ is a convex subset of the plane (β, h). Since $\mathcal{D}$ is the upper graph of $\beta \mapsto \mathsf{h}_c(\beta)$, the critical curve $\mathsf{h}_c(\cdot)$ of the general model (2.10) is a convex function, hence it is continuous (as we show below that it is finite everywhere).

We define the critical lines h_c^{pin} and h_c^{cop} of the pinning and copolymer models in analogy with (3.1), in terms of $\textsc{f}^{\text{pin}}$ and $\textsc{f}^{\text{cop}}$ respectively. Recalling (2.16), it follows that

$$h_c^{\text{pin}}(\cdot) = \mathsf{h}_c^{\psi^{\text{pin}}}(\cdot)\,, \qquad h_c^{\text{cop}}(\beta) = \frac{1}{2\beta}\mathsf{h}_c^{\psi^{\text{cop}}}(-2\beta) \tag{3.2}$$

In particular, the critical curve of the pinning model $h_c^{\text{pin}}(\cdot)$ is convex. This is not true in general for the copolymer, as $h_c^{\text{cop}}(\cdot)$ is the ratio between a convex function and a linear function, which in not necessarily convex.

3.2. **Easy bounds on the critical curve.** Let us prove some basic bounds on the critical curve $\mathsf{h}_c(\cdot)$. Recall that $\mathsf{h}_c(0) = 0$, and since $\beta \mapsto \mathsf{F}(\beta, h)$ is increasing, we have $\mathsf{F}(\beta, h) \geq \mathsf{F}(0, h)$, hence we obtain the lower bound $\mathsf{h}_c(\beta) \geq 0$ for every $\beta \geq 0$. On the other hand, applying Jensen's inequality we get the important *annealed (upper) bound*: since $\psi^2 = \psi$, because $\psi \in \{0, 1\}$ both for $\psi = \psi^{\mathrm{pin}}$ and $\psi = \psi^{\mathrm{cop}}$, we have

$$\begin{aligned} \mathsf{F}(\beta, h) \, &\leq \, \lim_{N\to\infty} \frac{1}{N} \, \log \mathbf{E}\,\mathbb{E}\big[e^{\sum_{n=1}^N (\beta\omega_n - h)\psi(S_{[n-1,n]})}\big] \\ &= \, \lim_{N\to\infty} \frac{1}{N} \, \log \mathbf{E}\big[e^{\sum_{n=1}^N (\log \mathrm{M}(\beta) - h)\psi(S_{[n-1,n]})}\big] \, = \, \mathsf{F}(0, h - \log \mathrm{M}(\beta))\,, \end{aligned}$$

which shows that $\mathsf{F}(\beta, h) = 0$ if $h - \log \mathrm{M}(\beta) \geq \mathsf{h}_c(0) = 0$. Summarizing:

$$\forall \beta \geq 0: \qquad 0 \, \leq \, \mathsf{h}_c(\beta) \, \leq \, \mathsf{h}_c^{\mathrm{ANN}}(\beta) \, := \, \log \mathrm{M}(\beta)\,. \tag{3.3}$$

It follows from (3.4) that $\mathsf{h}_c(\beta) < +\infty$ for every $\beta \geq 0$, which together with convexity implies that the critical curve $\beta \mapsto \mathsf{h}_c(\beta)$ is *continuous and strictly increasing.*

Recalling (3.2), the bounds in (3.3) are directly applicable to the pinning and copolymer models, provided in the copolymer case we properly define the annealed curve:

$$h_c^{\mathrm{ANN,pin}}(\beta) \, := \, \log \mathrm{M}(\beta)\,, \qquad h_c^{\mathrm{ANN,cop}}(\beta) \, := \, \frac{\log \mathrm{M}(-2\beta)}{2\beta}\,. \tag{3.4}$$

Note that for small $\beta > 0$ the curve $h_c^{\mathrm{ANN,pin}}(\beta) = \frac{1}{2}\beta^2 + o(\beta^2)$ is quadratic, whereas $h_c^{\mathrm{ANN,cop}}(\beta) = \beta + o(\beta)$ is linear. This is only due to the different parametrization.

3.3. **Improved bounds on the critical curve.** The lower bound in (3.3) can be substantially improved. Let us introduce the following rescaling of the annealed critical curve:

$$\mathsf{h}^{\mathrm{MBG}}(\beta) \, := \, \big(\tfrac{1}{1+\alpha}\big)^{-1} \log \mathrm{M}\big(\tfrac{\beta}{1+\alpha}\big) \, = \, \big(\tfrac{1}{1+\alpha}\big)^{-1} \mathsf{h}_c^{\mathrm{ANN}}\big(\tfrac{1}{1+\alpha}\beta\big)\,, \quad \forall \beta \geq 0\,. \tag{3.5}$$

Then we can improve (3.4) to

$$\mathsf{h}^{\mathrm{MBG}}(\beta) \, \leq \, \mathsf{h}_c(\beta) \, \leq \, \mathsf{h}_c^{\mathrm{ANN}}(\beta)\,, \qquad \forall \beta \in [0, \infty)\,. \tag{3.6}$$

The acronym "MBG" stands for Monthus — who introduced the curve in the physics literature as a candidate for the true critical curve of the copolymer model, cf. [23] — and Bodineau-Giacomin — who gave a rigorous proof of the lower bound in (3.6) for copolymer models, cf. [8]. Their original proof will be presented in §5.

Recalling (3.2), the bounds in (3.6) holds also for the critical lines h_c^{pin} and h_c^{cop} of the pinning and copolymer models: for every $\beta \in [0, \infty)$

$$h^{\mathrm{MBG,pin}}(\beta) \, \leq \, h_c^{\mathrm{pin}}(\beta) \, \leq \, h_c^{\mathrm{ANN,pin}}(\beta)\,, \quad h^{\mathrm{MBG,cop}}(\beta) \, \leq \, h_c^{\mathrm{cop}}(\beta) \, \leq \, h_c^{\mathrm{ANN,cop}}(\beta)\,, \tag{3.7}$$

provided in the copolymer model we define properly the MBG curve:

$$\begin{aligned} h^{\mathrm{MBG,pin}}(\beta) \, &:= \, \big(\tfrac{1}{1+\alpha}\big)^{-1} \log \mathrm{M}\big(\tfrac{\beta}{1+\alpha}\big) \, = \, \big(\tfrac{1}{1+\alpha}\big)^{-1} h_c^{\mathrm{ANN,pin}}\big(\tfrac{1}{1+\alpha}\beta\big) \\ h^{\mathrm{MBG,cop}}(\beta) \, &:= \, \big(\tfrac{2\beta}{1+\alpha}\big)^{-1} \log \mathrm{M}\big(\tfrac{-2\beta}{1+\alpha}\big) \, = \, h_c^{\mathrm{ANN,cop}}\big(\tfrac{1}{1+\alpha}\beta\big)\,. \end{aligned} \tag{3.8}$$

The proof of the lower bound (3.7) in the pinning case is given in cf. §10.2, but we stress from now that it requires more sophisticated techniques with respect to the copolymer case, a fact which is a priori surprising. The same techniques allow to show that for copolymer models the lower bound in (3.7) is actually strict as soon as $\beta > 0$, cf. 10.3.

Remark 3.1. It was first proved in [2], in a very general setting, that $h_c^{\mathrm{pin}}(\beta) > 0$ for $\beta > 0$, that is the pinning model with $\beta > 0$ and $h = 0$ is localized (even if on average the disorder is centered). A quantitative improvement of this result, in a slightly less general setting, is provided by [18, Theorem 5.2]. The MBG lower bound in (3.7) improves this estimates through an explicit quadratic lower bound on $h_c(\beta)$, which is of the same order as the annealed upper bound (the first quadratic lower bound was obtained in [24], for a simple random walk model of polymer plus pinning). We point out that, in the copolymer case, the fact that $h_c^{\mathrm{cop}}(\beta) > 0$ for every $\beta > 0$ has been known for a long time, cf. [26, 10].

Developing the upper and lower bounds in (3.7) as $\beta \downarrow 0$, we obtain the estimates

$$\frac{1}{1+\alpha}\,\frac{1}{2} \le \frac{h_c^{\mathrm{pin}}(\beta)}{\beta^2} + o(1) \le \frac{1}{2}\,, \qquad \frac{1}{1+\alpha} \le \frac{h_c^{\mathrm{cop}}(\beta)}{\beta} + o(1) \le 1\,, \quad \text{as } \beta \downarrow 0\,. \tag{3.9}$$

We stress that the differences between the pinning and copolymer case, namely the order β^2 instead of β and the constant $\frac{1}{2}$ instead of 1, are simply due to the different parametrization of these models, cf. (2.11). The asymptotic bounds in (3.9) can be improved.

- For the copolymer model, it was shown in [28] that the upper bound can be made strictly smaller than 1 (with an estimate depending only on α). Recently it has been shown in [11] that the lower bound can be made strictly larger than $\frac{1}{1+\alpha}$, with an explicit expression for $\alpha > 1$. We present this proof in §10.4. Previous improvements were obtained in [9]. For a numerical study of the simple random walk case, corresponding to $\alpha = \frac{1}{2}$, we refer to [13].
- For the pinning model, the lower bound can be improved to $\frac{1}{2}$, matching asymptotically the upper bound, for $\alpha \in (0, 1)$; as a matter of fact, for $\alpha \in (0, \frac{1}{2})$ the critical curve $h_c(\beta)$ actually *coincides* with the annealed bound $h_c^{\mathrm{ANN}}(\beta)$ for β small, see section 11. On the other hand, for $\alpha > 1$ the upper bound can be made strictly smaller than $\frac{1}{2}$, cf. [19, Theorem 6.1], therefore the lower bound provided by (3.7) provides an explicit and non-trivial estimate.

3.4. **The weak coupling limit of the copolymer model.** The behavior of $h_c^{\mathrm{cop}}(\beta)$ in the weak disorder limit $\beta \to 0$, is a very interesting object (universality, cf. [12]). **TO BE COMPLETED**.

3.5. **Relevance and irrelevance of disorder for pinning models. TO BE COMPLETED**.

4. The rare stretch strategy

We introduce in this section a simple and powerful technique, known as *rare stretch strategy*, which is cornerstone of all the proofs that will be developed in this notes. This was originally introduced in [8], to prove the MBG lower bound (3.7) for the copolymer model. We present this proof as a first application of the rare stretch strategy in §5.

4.1. **The $(\mathcal{G}, \mathcal{C})$-rare stretch strategy.** For every $\ell \in \mathbb{N}$ let $\mathcal{A}_\ell \subseteq \mathbb{R}^\ell$ be a set of "good atypical stretches", such that there exist $\mathcal{G} \in [0, \infty)$, $\mathcal{C} \in [0, \infty)$ such that for a diverging sequence of $\ell \in \mathrm{T}\mathbb{N}$ (we recall that T is the period of the underlying walk):

- for all $\omega = \omega_{[1,\ell]} := (\omega_1, \ldots, \omega_\ell) \in \mathcal{A}_\ell$ we have $\frac{1}{\ell} \log \mathsf{Z}^{\mathrm{c}}_{\ell,\omega,\beta,h} \ge \mathcal{G}$;

- we have $\frac{1}{\ell}\log\mathbb{P}(\mathcal{A}_\ell) \geq -\mathcal{C}$.

We call this an $(\mathcal{G},\mathcal{C})$-strategy, where $\mathcal{G}$ stands for *gain* and $\mathcal{C}$ for *cost*. We claim that

$$\mathsf{F}(\beta,h) > 0 \quad \text{(that is } (\beta,h)\in\mathcal{L}\text{, that is } h<\mathsf{h}_c(\beta)) \qquad \text{if} \qquad \mathcal{G}-(1+\alpha)\mathcal{C} > 0\,. \tag{4.1}$$

The idea is that the stretches $\omega_{[1,\ell]} \in \mathcal{A}_\ell$ may be worth visiting, because the polymer gains $\mathsf{Z}^{\mathsf{c}}_{\ell,\omega_{[1,\ell]},\beta,h} \geq e^{\ell\mathcal{G}}$ and at the same time they are not too unlikely, because $\mathbb{P}(\mathcal{A}_\ell) \geq e^{-\ell\mathcal{C}}$. Note that the condition is on the *constrained* partition function.

Fix $\ell \in \mathrm{T}\mathbb{N}$ large enough, so that the above conditions hold. For $\omega\in\mathbb{R}^{\mathbb{N}}$, we denote by $T_1(\omega), T_2(\omega),\ldots$ the instants in $\ell\mathbb{N}$ that are endpoints of good stretches:

$$T_1(\omega) := \inf\{N\in\ell\mathbb{N}:\ \omega_{(N-\ell,\ell]}\in\mathcal{A}_\ell\}\,, \qquad T_k(\omega) := T_1(\vartheta^{T_1(\omega)+\ldots+T_{k-1}(\omega)}(\omega))\,.$$

Note that $\{T_k\}_{k\in\mathbb{N}}$ are i.i.d. with marginal laws given by $\ell\cdot\mathrm{Ge}(\mathbb{P}(\mathcal{A}_\ell))$. In particular,

$$\mathbb{E}(T_1) = \ell/\mathbb{P}(\mathcal{A}_\ell)\,.$$

Observe that

$$\mathsf{Z}_{T_1+\ldots+T_k,\omega} \geq \prod_{i=1}^k K(T_i-\ell)\,\mathsf{Z}_{\ell,\vartheta^{(T_1+\ldots+T_{i-1})-\ell}\omega} \geq e^{k\mathcal{G}\ell}\prod_{i=1}^k K(T_i-\ell)\,,$$

where we set $K(0):=1$ for convenience. Therefore by the strong law of large numbers

$$\begin{aligned}\mathsf{F}(\beta,h) &= \lim_{N\to\infty}\frac{1}{N}\log\mathsf{Z}_{N,\omega,\beta,h} = \lim_{k\to\infty}\frac{1}{T_1+\ldots+T_k}\log\mathsf{Z}_{T_1+\ldots+T_k,\omega,\beta,h}\\ &\geq \frac{1}{\mathbb{E}(T_1(\omega))}\big\{\ell\mathcal{G}+\mathbb{E}(\log K(T_1(\omega)-\ell))\big\}\,.\end{aligned}$$

By (2.3), for every $\varepsilon>0$ there exists $c'>0$ such that $K(n)\geq c'/(n+1)^{1+\alpha+\varepsilon}$ for all $n\in\mathrm{T}\mathbb{N}_0$ (we write $(n+1)$ in order to include also $n=0$). Since $\ell\geq 1$, it follows that $K(n)\geq c'/(n+\ell)^{1+\alpha+\varepsilon}$ — just for ease of computation — and by Jensen's inequality

$$\begin{aligned}\mathsf{F}(\beta,h) &\geq \frac{1}{\mathbb{E}(T_1)}\big\{\ell\mathcal{G}-(1+\alpha+\varepsilon)\log\mathbb{E}(T_1)+\log c'\big\}\\ &= \frac{\mathbb{P}(\mathcal{A}_\ell)}{\ell}\big\{\ell\mathcal{G}+(1+\alpha+\varepsilon)\log\mathbb{P}(\mathcal{A}_\ell)-(1+\alpha)\log\ell+\log c'\big\}\\ &= e^{-I\ell}\bigg\{\big(\mathcal{G}-(1+\alpha)\mathcal{C}\big)-\varepsilon\mathcal{C}-(1+\alpha+\varepsilon)\frac{\log\ell}{\ell}+\frac{\log c'}{\ell}\bigg\}\,.\end{aligned}$$

If $\mathcal{G}-(1+\alpha)\mathcal{C}>0$, it is clear that we can choose $\varepsilon>0$ small enough and $\ell\in\mathbb{N}$ large enough (but finite!) so that the right hand side is strictly positive. This proves (4.1).

Remark 4.1. It is worth stressing that all the preceding argument still work for a general ergodic environment $(\{\omega_n\}_{n\in\mathbb{N}},\mathbb{P})$, beyond the i.i.d. case: it suffices to use the ergodic theorem instead of the strong law of large numbers. However, in the ergodic case it may be difficult to check the condition $\frac{1}{\ell}\log\mathbb{P}(\mathcal{A}_\ell)\geq-\mathcal{C}$.

5. The MBG lower bound for copolymers

A first application of the rare stretch strategy — indeed, the original one [8] — is the derivation of the Monthus-Bodineau-Giacomin lower bound on the critical curve $h_c^{\mathrm{cop}}(\cdot)$ of the copolymer model, cf. (3.7), that we rewrite for convenience:

$$h_c^{\mathrm{cop}}(\beta) \,\geq\, h^{\mathrm{MBG,cop}}(\beta)\,, \qquad \forall \beta \geq 0\,, \tag{5.1}$$

where

$$h^{\mathrm{MBG,cop}}(\beta) \,:=\, \big(\tfrac{2\beta}{1+\alpha}\big)^{-1} \log \mathrm{M}\big(\tfrac{-2\beta}{1+\alpha}\big) \,=\, h_c^{\mathrm{ANN,cop}}\big(\tfrac{1}{1+\alpha}\beta\big)\,. \tag{5.2}$$

The proof of the analogous bound in the pinning case is less direct, cf. §10.2.

We need the following lemma, that will be used also later.

Lemma 5.1 (Entropy inequality). *Given two laws μ, ν on the same measurable space with $\nu \ll \mu$ (ν is absolutely continuous with respect to μ), for every event A we have*

$$\mu(A) \,\geq\, \nu(A) \exp\bigg(- \frac{1}{\nu(A)} \big(h(\nu|\mu) + e^{-1} \big) \bigg)\,, \tag{5.3}$$

where the relative entropy *$h(\nu|\mu)$ is defined by*

$$h(\nu|\mu) \,:=\, \mathrm{E}_\nu \bigg(\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \bigg) \,=\, \mathrm{E}_\mu \bigg(\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \log \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \bigg)\,. \tag{5.4}$$

Proof. Assume first that also $\mu \ll \nu$. Then by Jensen

$$\mu(A) \,=\, \mathrm{E}_\nu \bigg(1_A \frac{\mathrm{d}\mu}{\mathrm{d}\nu} \bigg) \,=\, \mathrm{P}_\nu(A) \cdot \mathrm{E}_\nu \big(e^{-\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu}} \big| A \big) \,\geq\, \mathrm{P}_\nu(A) \cdot e^{-\mathrm{E}_\nu (\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu} | A)}\,,$$

and since $x \log x \geq e^{-1}$ for every $x \geq 0$, we obtain (5.3):

$$\mathrm{E}_\nu \bigg(\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \bigg| A \bigg) \,=\, \frac{1}{\nu(A)}\, \mathrm{E}_\mu \bigg(1_A \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \log \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \bigg) \,\geq\, \frac{1}{\nu(A)} \bigg(\mathrm{E}_\mu \bigg(\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \log \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \bigg) - e^{-1} \bigg)\,.$$

To remove the extra-assumption $\mu \ll \nu$, we apply the inequality to $\nu_\varepsilon := (1-\varepsilon)\nu + \varepsilon\mu$ and we show that $h(\nu_\varepsilon|\mu) \to h(\nu|\mu)$ as $\varepsilon \downarrow 0$. Setting $X := \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$, we have $h(\nu_\varepsilon|\mu) = \mathrm{E}_\mu(X_\varepsilon \log X_\varepsilon)$, where $X_\varepsilon := (1-\varepsilon)X + \varepsilon$. Since $x \mapsto x \log x$ is a convex function, we have

$$X_\varepsilon \log X_\varepsilon \,\leq\, (1-\varepsilon) X \log X + \varepsilon 1 \log 1 \,=\, (1-\varepsilon) X \log X\,,$$

hence $h(\nu_\varepsilon|\mu) = \mathrm{E}_\mu(X_\varepsilon \log X_\varepsilon) \to \mathrm{E}_\mu(X \log X) = h(\nu|\mu)$ by dominated convergence (observe that $X_\varepsilon \log X_\varepsilon \geq \inf_{x \geq 0}(x \log x) \geq e^{-1}$). □

Let us now come to the the proof of (5.1). We recall that for every $\ell \in \mathrm{T}\mathbb{N}$

$$Z^{\mathrm{cop,c}}_{\ell,\omega,\beta,h} \,=\, \mathbf{E}\big[e^{-2\beta \sum_{n=1}^{\ell} (\omega_n + h) 1_{\{S_{n-1} \leq 0,\, S_n \leq 0\}}}\, 1_{\{S_\ell = 0\}} \big]\,.$$

We fix $q \in \mathbb{R}$ (think of $q < -h$) and consider the following set of good atypical stretches:

$$\mathcal{A}_\ell \,:=\, \bigg\{ \omega = \omega_{[1,\ell]} \in \mathbb{R}^\ell : \, \frac{1}{\ell} \sum_{i=1}^{\ell} \omega_i \,\leq\, q \bigg\}\,.$$

Let us fix an arbitrary $\varepsilon > 0$. We claim that $\mathcal{A}_\ell$ defines a $(\mathcal{G}, \mathcal{C})$-strategy with

$$\mathcal{G} \,=\, -2\beta(q+h) - \varepsilon\,, \qquad \mathcal{C} \,=\, (\delta (\log \mathrm{M})'(\delta) - \log \mathrm{M}(\delta)) + \varepsilon\,. \tag{5.5}$$

The identification of $\mathcal{G}$ is easy: restricting the expectation to trajectories that stay in the lower half-plane, for $\omega \in \mathcal{A}_\ell$ we obtain the lower bound

$$\begin{aligned} Z^{\mathrm{cop,c}}_{\ell,\omega,\beta,h} &\ge \mathbf{E}\big[e^{-2\beta\sum_{n=1}^{\ell}(\omega_n+h)1_{\{S_{n-1}\le 0,\, S_n\le 0\}}}\, 1_{\{S_1\le 0,\dots,S_{\ell-1}\le 0, S_\ell=0\}}\big] \\ &= e^{-2\beta(\sum_{n=1}^{\ell}\omega_n)-2\beta h\ell}\,\mathbf{P}(S_1\le 0,\dots,S_{\ell-1}\le 0, S_\ell=0) \\ &\ge \exp\left\{\ell\bigg(-2\beta q-2\beta h+\frac{1}{\ell}\big(\log K(\ell)-\log 2\big)\bigg)\right\}, \end{aligned}$$

having applied (2.6). Since $\frac{1}{\ell}(\log K(\ell)-\log 2)\to 0$ as $\ell\to\infty$, by (2.2), for large $\ell \in \mathrm{T}\mathbb{N}$ we have indeed $\frac{1}{\ell}\log Z^{\mathrm{cop,c}}_{\ell,\omega,\beta,h} \ge \mathcal{G}$ for all $\omega\in\mathcal{A}_\ell$, with the $\mathcal{G}$ given in (5.5), as required.

It remains to determine $\mathcal{C}$, that is to estimate from below the rate of exponential decay of $\mathbb{P}(\mathcal{A}_\ell)$. This is one of the most basic large deviations result (Cramer's theorem), that we obtain as an application of the entropy inequality (5.3).

For $\delta\in\mathbb{R}$, we introduce the *tilted law* $\widetilde{\mathbb{P}}_\delta$, under which $\{\omega_n\}_{n\in\mathbb{N}}$ are i.i.d. with

$$\widetilde{\mathbb{P}}_\delta(\omega_1\in \mathrm{d}x) := e^{\delta\omega_1-\log \mathrm{M}(\delta)}\,\mathbb{P}(\omega_1\in\mathrm{d}x)\,. \tag{5.6}$$

Note that in the Gaussian case we have $\omega_n \sim N(\delta,1)$ under $\widetilde{\mathbb{P}}_\delta$. We also set $\mathrm{M}_\delta(t) := \widetilde{\mathbb{E}}_\delta(e^{t\omega_1}) = \mathrm{M}(t+\delta)/\mathrm{M}(\delta)$ and

$$m_\delta := \widetilde{\mathbb{E}}_\delta(\omega_1) = (\log \mathrm{M})'(\delta) = \frac{\mathrm{M}'(\delta)}{\mathrm{M}(\delta)} = \delta+o(\delta)\,, \quad \text{as } \delta\to 0\,. \tag{5.7}$$

The law $\widetilde{\mathbb{P}}_\delta$ restricted to $(\omega_1,\dots,\omega_\ell)$ is absolutely continuous with respect to $\mathbb{P}$, with Radon-Nikodym density given by

$$\frac{\mathrm{d}\widetilde{\mathbb{P}}_\delta|_\ell}{\mathrm{d}\mathbb{P}|_\ell}(\omega_1,\dots,\omega_\ell) = e^{\sum_{i=1}^{\ell}(\delta\omega_i-\log\mathrm{M}(\delta))}\,,$$

hence the corresponding relative entropy is given by

$$h(\widetilde{\mathbb{P}}_\delta|_\ell\,|\,\mathbb{P}|_\ell) = \widetilde{\mathbb{E}}_\delta\bigg(\log\frac{\mathrm{d}\widetilde{\mathbb{P}}_\delta|_\ell}{\mathrm{d}\mathbb{P}|_\ell}(\omega_1,\dots,\omega_\ell)\bigg) = \ell\big(\delta\, m_\delta-\log\mathrm{M}(\delta)\big)\,. \tag{5.8}$$

By the entropy inequality (5.3) we get

$$\mathbb{P}(\mathcal{A}_\ell) \ge \widetilde{\mathbb{P}}_\delta(\mathcal{A}_\ell)\,\exp\bigg(-\frac{1}{\widetilde{\mathbb{P}}_\delta(\mathcal{A}_\ell)}\Big(h(\widetilde{\mathbb{P}}_\delta|_\ell\,|\,\mathbb{P}|_\ell)+\frac{1}{e}\Big)\bigg)\,.$$

Now note that, if $\delta\in\mathbb{R}$ is such that $m_\delta<q$, the event $\mathcal{A}_\ell$ is $\widetilde{\mathbb{P}}_\delta$-typical: in fact, by the weak law of large numbers

$$\lim_{\ell\to\infty}\widetilde{\mathbb{P}}_\delta(\mathcal{A}_\ell) = 1\,,$$

therefore for such values of δ and for large ℓ we have $\frac{1}{\ell}\log\mathbb{P}(\mathcal{A}_\ell)\ge -\mathcal{C}$, with $\mathcal{C}$ given in (5.5), as required.

We can complete the proof of (5.1). Having given a $(\mathcal{G},\mathcal{C})$ strategy, it follows that $(\beta,h)\in\mathcal{L}$ — equivalently, $h<h_c^{\mathrm{cop}}(\beta)$ — if $\mathcal{G}-(1+\alpha)\mathcal{C}>0$, that is, recalling (5.5),

$$-2\beta(q+h)-(1+\alpha)(\delta\, m_\delta-\log\mathrm{M}(\delta))-2\varepsilon > 0\,.$$

Note that we are free to choose $q\in\mathbb{R}$, $\varepsilon>0$ and $\delta\in\mathbb{R}$ as long as $q>m_\delta$. Choosing $\delta=\frac{-2\beta}{1+\alpha}$, we can rewrite this condition as

$$-2\varepsilon - 2\beta(q-m_\delta) + 2\beta\big\{\tfrac{1+\alpha}{2\beta}\log\mathrm{M}\big(\tfrac{-2\beta}{1+\alpha}\big)-h\big\} > 0\,.$$

The first two terms in the right hand side can be made as close to zero as we wish, by choosing $\varepsilon > 0$ small enough and q close enough to m_δ. Therefore we have localization, that is $h < h_c^{\mathrm{cop}}(\beta)$, if the term in brackets is strictly positive. This shows that (5.1) must hold.

6. Smoothing effect of quenched disorder

As a second application of the rare stretch strategy, cf. §4.1, we present the beautiful proof of Giacomin and Toninelli [20, 21] that, as soon as disorder is present, the phase transition for pinning and copolymer models is always at least of second order.

Recall the definition of the tilted law $\widetilde{\mathbb{P}}_\delta$, cf. (5.6) and the neighboring lines. Let us set

$$\mathsf{F}_N(\beta, h; \delta) := \frac{1}{N}\widetilde{\mathbb{E}}_\delta\big[\log \mathsf{Z}_{N,\omega,\beta,h}\big] = \frac{1}{N}\mathbb{E}\big[(\log \mathsf{Z}_{N,\omega,\beta,h})\, e^{\sum_{n=1}^N \delta\omega_n - \log \mathrm{M}(\delta)}\big]\,, \tag{6.1}$$

and then define the generalized free energy:

$$\mathsf{F}(\beta, h; \delta) := \lim_{N\to\infty} \mathsf{F}_N(\beta, h; \delta) = \lim_{\substack{N\to\infty\\ N\in\tau\mathbb{N}}} \frac{1}{N}\log \mathsf{Z}^{\mathrm{c}}_{N,\omega,\beta,h}\,, \quad \widetilde{\mathbb{P}}_\delta\text{-a.s. and in } L^1(\widetilde{\mathbb{P}}_\delta)\,, \tag{6.2}$$

so that $\mathsf{F}(\beta, h) = \mathsf{F}(\beta, h; 0)$. Note that the a.s. convergence in (6.2) has been stated for the constrained partition function $\mathsf{Z}^{\mathrm{c}}_{N,\omega,\beta,h}$, because this is what we are going to need, but of course it holds also for the free one $\mathsf{Z}_{N,\omega,\beta,h}$.

We first prove a smoothing inequality in δ for the tilted free energy $\mathsf{F}(\beta, \mathsf{h}_c(\beta); \delta)$.

Theorem 6.1. *Assume that $M(t) < \infty$ for every $t \in (-t_0, t_0)$, where $t_0 > 0$. The following relation holds for every $\beta \geq 0$ and $0 \leq \delta < t_0$:*

$$0 \leq \mathsf{F}(\beta, \mathsf{h}_c(\beta); \delta) \leq (1+\alpha)\big(\delta\,(\log \mathrm{M})'(\delta) - \log \mathrm{M}(\delta)\big) = (1+\alpha)\frac{\delta^2}{2} + O(\delta^3) \quad \text{as } \delta \downarrow 0\,.$$

Proof. We are going to apply a $(\mathcal{G}, \mathcal{C})$ rare stretch strategy. This may appear strange, since we are after an *upper* bound, while the rare stretch strategy is based on a *lower* bound. Everything will be clear at the end.

Fix $\beta > 0$, $h \in \mathbb{R}$ and a small $\varepsilon > 0$, and define the set of good atypical stretches as

$$\mathcal{A}_\ell := \left\{(\omega_1, \ldots, \omega_\ell) \in \mathbb{R}^\ell : \frac{1}{\ell}\log \mathsf{Z}^{\mathrm{c}}_{\ell,\omega,\beta,h} \geq \mathsf{F}(\beta, h; \delta) - \varepsilon\right\},$$

so that $\mathcal{G} = \mathsf{F}(\beta, h; \delta) - \varepsilon$. It remains to determine $\mathcal{C}$, for which we need to estimate the probability of $\mathbb{P}(\mathcal{A}_\ell)$ from below. Note that, by (6.2), the event $\mathcal{A}_\ell$ is typical for $\widetilde{\mathbb{P}}_\delta$:

$$\lim_{\ell\to+\infty} \widetilde{\mathbb{P}}_\delta(\mathcal{A}_\ell) = 1\,. \tag{6.3}$$

Recalling the entropy inequality (5.3):

$$\mathbb{P}(\mathcal{A}_\ell) \geq \widetilde{\mathbb{P}}_\delta(\mathcal{A}_\ell)\, \exp\left(-\frac{1}{\widetilde{\mathbb{P}}_\delta(\mathcal{A}_\ell)}\big(h(\widetilde{\mathbb{P}}_\delta|_\ell\,|\,\mathbb{P}|_\ell) + e^{-1}\big)\right),$$

it follows by (5.8) that for large ℓ

$$\frac{1}{\ell}\log \mathbb{P}(\mathcal{A}_\ell) \geq -\big(\delta\, m_\delta - \log \mathrm{M}(\delta)\big) - \varepsilon =: -\mathcal{C}\,,$$

where we remind that $m_\delta = (\log \mathrm{M})'(\delta) = \widetilde{\mathbb{E}}_\delta(\omega_1)$.

Because of our $(\mathcal{G}, \mathcal{C})$-strategy, we know by (4.1) that $\mathsf{F}(\beta, h) > 0$ if

$$\mathcal{G} - (1+\alpha)\mathcal{C} = \mathsf{F}(\beta, h; \delta) - (1+\alpha)\big(\delta\, m_\delta - \log \mathrm{M}(\delta)\big) - 2\varepsilon > 0\,.$$

Now comes the brilliant idea. Let us choose $h = \mathsf{h}_c(\beta)$, so that $\mathsf{F}(\beta, h) = 0$. It follows that $\mathcal{G} - (1+\alpha)\mathcal{C} \le 0$, that is

$$\mathsf{F}(\beta, h; \delta) \;\le\; (1+\alpha)\big(\delta\, m_\delta - \log \mathrm{M}(\delta)\big) \;+\; 2\varepsilon\,.$$

Since this equality holds for every $\varepsilon > 0$, it must hold also for $\varepsilon = 0$. □

Now observe that, if ω_n are i.i.d. $N(0,1)$ Gaussian variables, we can write $\mathsf{F}^{[a,b]}(\beta, h; \delta) = \mathsf{F}(\beta, h-\delta; 0)$, therefore Theorem 6.1 gives indeed information about the regularity of the phase transition. **TO BE COMPLETED**.

The interesting fact is that $\mathsf{F}^{[a,b]}(\beta, h; \delta)$ dominates $\mathsf{F}^{[a,b]}(\beta, h-\delta; 0)$ for general disorder distributions, hence the smoothing effect of disorder is general. More precisely, with the same techniques as in [20] one can prove the following result.

Theorem 6.2. *Assume that $M(t) < \infty$ for every $t \in (-t_0, t_0)$, where $t_0 > 0$. For every $\beta > 0$ there exist $C = C(\beta) > 0$ and $\delta_0 = \delta_0(\beta) > 0$ such that*

$$\textsc{f}(\beta, h; \delta) \;\ge\; \textsc{f}\big(\beta, h - C(\beta)\,\beta\,\delta\big)\,, \qquad \forall 0 \le \delta \le \delta_0\,. \tag{6.4}$$

The constants $C(\beta)$ and $\delta_0(\beta)$ can actually be taken independent of β, as long as β ranges in a bounded interval $[0, M]$.

Corollary 6.3. *Assume that $M(t) < \infty$ for every $t \in (-t_0, t_0)$, where $t_0 > 0$. For every $\beta > 0$ there exist $C' = C'(\beta) > 0$ and $\delta_0' = \delta_0'(\beta) > 0$ such that*

$$0 \;\le\; \textsc{f}(\beta, \mathsf{h}_c(\beta) - \delta) \;\le\; C'\,\delta^2\,, \qquad \forall 0 \le \delta \le \delta_0'\,. \tag{6.5}$$

7. Some entropy background

7.1. **Absolute and relative entropy.** We have already encountered the notion of *relative entropy*: given any two probabilities ν, μ on the same measurable space, one defines

$$h(\nu|\mu) \;:=\; \begin{cases} \mathrm{E}_\nu\left(\log \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right) \;=\; \mathrm{E}_\mu\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu} \log \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right) & \text{if } \nu \ll \mu \\ +\infty & \text{otherwise} \end{cases}\,. \tag{7.1}$$

Let us set $f(x) := x\,\log x$ for $x \ge 0$, with $0 \log 0 := 0$. Since f is convex, for $\nu \ll \mu$ we can write, by Jensen,

$$h(\nu|\mu) \;=\; \mathrm{E}_\mu(f(\tfrac{\mathrm{d}\nu}{\mathrm{d}\mu})) \;\ge\; f(\mathrm{E}_\mu(\tfrac{\mathrm{d}\nu}{\mathrm{d}\mu})) \;=\; 0\,,$$

and the equality holds if and only if $\frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ is constant μ-a.s., that is $\nu = \mu$. Therefore, one can view $h(\nu|\mu)$ as a sort of measure of how ν is different from μ.[†]

For the recent development in [6, 7, 11], a central role is played by the concept of *absolute entropy*, that can be introduced as follows. Given a law μ on a measurable space $(F, \mathcal{F})$, we define two laws on the product space $(F \times F, \mathcal{F} \otimes \mathcal{F})$ (equipped with the product σ-field).

- The diagonal law $\mu_{2,\mathrm{d}}$ is defined as the joint law of the couple (X, X), where X is a random variable with law μ. More explicitly, for every $C \in \mathcal{F} \otimes \mathcal{F}$ we have $\mu_{2,\mathrm{d}}(C) := \mu\{x \in F : \; (x,x) \in C\}$.
- The product law $\mu^{\otimes 2} = \mu \otimes \mu$ is defined as the joint law of the couple (X, Y), where X and Y are *independent* random variables with law μ. More explicitly, on product sets $A \times B \in \mathcal{F} \otimes \mathcal{F}$, with $A, B \in \mathcal{F}$, we have $\mu^{\otimes 2}(A \times B) := \mu(A)\mu(B)$.

[†]Note however that $h(\nu|\mu)$ is *not* a distance between probabilities (it is not symmetric, it does not satisfy the triangle inequality).

The absolute entropy of μ is defined as the relative entropy of $\mu_{2,\mathrm{d}}$ with respect of $\mu^{\otimes 2}$:

$$h(\mu) := h(\mu_{2,\mathrm{d}}|\mu^{\otimes 2}) \in [0,+\infty] . \tag{7.2}$$

Note that the diagonal law $\mu_{2,\mathrm{d}}$ is concentrated on the diagonal $D = \{(x,y) \in F \times F : x = y\} \subseteq F \times F$. Therefore, for most non-discrete laws μ, the law $\mu_{2,\mathrm{d}}$ is *not* absolutely continuous with respect to the product law $\mu^{\otimes 2}$, in which case $h(\mu) = \infty$.

On the other hand, for discrete laws $\mu = \sum_{n\in\mathbb{N}} p_n\, \delta_{x_n}$, where $\{x_n\}_{n\in\mathbb{N}}$ are pairwise *distinct* points in F and $\{p_n\}_{n\in\mathbb{N}}$ are non-negative numbers such that $\sum_{n\in\mathbb{N}} p_n = 1$, we have

$$\mu_{2,\mathrm{d}} = \sum_{m,n\in\mathbb{N}} p_m\, \mathbf{1}_{\{m=n\}}\, \delta_{(x_m,x_n)} , \qquad \mu^{\otimes 2} = \sum_{m,n\in\mathbb{N}} p_m\, p_n\, \delta_{(x_m,x_n)} ,$$

hence

$$\frac{\mathrm{d}\mu_{2,\mathrm{d}}}{\mu^{\otimes 2}}(x_n,x_m) = \frac{p_n \mathbf{1}_{\{n=m\}}}{p_n\, p_m} = \frac{1}{p_m}\, \mathbf{1}_{\{n=m\}} ,$$

and we obtain the basic formula (note the minus sign)

$$h(\mu) = -\sum_{n\in\mathbb{N}} p_n \log p_n \in [0,+\infty] . \tag{7.3}$$

In particular, if X is a random variable with law μ, then

$$h(\mu) = -\,\mathrm{E}\big[\log \mathrm{P}(X = x)|_{x=X}\big] . \tag{7.4}$$

Relation (7.3), or equivalently (7.4), is the way absolute entropy is usually defined for discrete laws $\mu = \sum_{n\in\mathbb{N}} p_n\, \delta_{x_n}$ (by the way, note that it only depends on the probabilities $\{p_n\}_{n\in\mathbb{N}}$ and not on the points $\{x_n\}_{n\in\mathbb{N}}$). However, relation (7.2) is nice for the intuition, because it shows that the larger $h(\mu)$, the more μ is "spread out".

For simplicity (and also necessity), henceforth we will work on finite or countable spaces.

7.2. **Entropy rate (absolute and relative).** Let F be a countable set, equipped with the discrete σ-field $\mathcal{F} = \mathcal{P}(F)$. The infinite product space $F^{\mathbb{N}}$ is equipped as usual with the product σ-field $\mathcal{F}^{\otimes\mathbb{N}}$ and we denote by $\{X_n\}_{n\in\mathbb{N}}$ the coordinate process, defined by $X_n(x) := x_n$ for for $x = (x_n)_{n\in\mathbb{N}} \in F^{\mathbb{N}}$. Note that a probability Q on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$ may be viewed as a joint law of the process $\{X_n\}_{n\in\mathbb{N}}$.

It turns out that the concept of relative entropy is not very useful for laws on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$. For instance, consider the case of i.i.d. laws Q, Q', that is $Q = q^{\otimes\mathbb{N}}$ and $Q' = (q')^{\otimes\mathbb{N}}$ where q, q' are laws on $(F, \mathcal{F})$. If $q \neq q'$, then Q is *not* absolutely continuous with respect to Q',[†] hence $h(Q|Q') = \infty$. In other words, for i.i.d. laws Q, Q', the relative entropy $h(Q|Q')$ is zero (obviously) if $Q = Q'$ and $+\infty$ as soon as $Q \neq Q'$.

For this reason, given two probabilities Q, Q' on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$, it is usually more useful to consider the *relative entropy rate* $H(Q|Q') \in [0,+\infty]$, defined in the following way. For every $n \in \mathbb{N}$, let Q_n denote the projection of Q on the first n components, i.e., the law of $(X_1,\ldots,X_n)$ on $(F^n, \mathcal{F}^{\otimes n})$ under Q, and define analogously Q'_n. Then we set

$$\begin{aligned} H(Q|Q') &:= \limsup_{n\to\infty} \frac{1}{n}\, h(Q_n|Q'_n) \\ &= \limsup_{n\to\infty} \frac{1}{n}\, \mathrm{E}_Q\left[\log \frac{Q((X_1,\ldots,X_n)=\xi)}{Q'((X_1,\ldots,X_n)=\xi)}\bigg|_{\xi=(X_1,\ldots,X_n)}\right] . \end{aligned} \tag{7.5}$$

[†] In fact, for $a \in F$, consider the set $A_a := \{x \in F^{\mathbb{N}} : \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{\{x_i=a\}} = q'(a)\} \in \mathcal{F}^{\otimes\mathbb{N}}$. By the strong law of large numbers, $Q'(A_a) = 1$ for every $a \in F$, while $Q(A_a) = 0$ if $q'(a) \neq q(a)$. Therefore, if $q \neq q'$, there exists $A \in \mathcal{F}^{\otimes\mathbb{N}}$ such that $Q'(A) = 1$ and $Q(A) = 0$, hence $Q \not\ll Q'$.

Analogously, the *absolute entropy rate* $H(Q) \in [0,\infty]$ of a law Q on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$ is defined as

$$\begin{aligned} H(Q) \,&:=\, \limsup_{n\to\infty} \frac{1}{n}\, h(Q_n) \\ &:=\, \limsup_{n\to\infty} \bigg(-\frac{1}{n}\, \mathrm{E}_Q \big[\log Q((X_1,\ldots,X_n)=\xi)|_{\xi=(X_1,\ldots,X_n)} \big] \bigg) . \end{aligned} \tag{7.6}$$

Remark 7.1. For i.i.d. laws $Q = q^{\otimes\mathbb{N}}$ we have $h(Q_n) = nh(Q_1)$, hence $H(Q) = h(Q_1) = h(q)$ and (7.6) holds as a limit. In general, however, the lim sup in both the above relation is not a true limit, unless some extra-assumption is made. As we are going to see, a key role is played by the notion of *ergodicity*.

7.3. **Ergodicity and the Asymptotic Equipartition Property.** We always assume that F is a countable set equipped with the discrete σ-field $\mathcal{F} = \mathcal{P}(F)$ (although several results do hold for in a much more general setting). The *(left) shift* operator $\vartheta : F^{\mathbb{N}} \to F^{\mathbb{N}}$ is defined by $(\vartheta x)_n := x_{n+1}$. A set $A \subseteq F^{\mathbb{N}}$ is said to be *shift-invariant* if $\vartheta^{-1}(A) = A$.[†]

- A law Q on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$ is said to be *stationary* if it is shift-invariant, i.e., $Q(A) = Q(\vartheta^{-1}(A))$ for all $A \in \mathcal{F}^{\otimes\mathbb{N}}$.
- A law Q on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$ is said to be *ergodic* it if is stationary and if $Q(A) \in \{0,1\}$ for every shift-invariant measurable event $A \in \mathcal{F}^{\otimes\mathbb{N}}$.

The set of ergodic laws on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$ will be denoted by $\mathcal{M}_1^{\mathrm{erg}}(F^{\mathbb{N}})$. We will also write "ergodic process" to mean "ergodic law".

Remark 7.2. Basic examples of ergodic laws are provided by i.i.d. laws. This is a consequence of Kolmogorov's 0–1 law, because every shift-invariant event is in the tail σ-field. Beyond i.i.d. laws, important examples of ergodic laws are given by irreducible and positive recurrent Markov chains started with the (unique) stationary law.

Let us stress, however, that ergodic processes can be strongly correlated. For an easy example, given two different points $a, b \in F$, consider the sequences $x = (a,b,a,b,\ldots)$ — i.e. $x_{2n} := a$ and $x_{2n-1} := b$ for every $n \in \mathbb{N}$ — and $y := \vartheta x = (b,a,b,a,\ldots)$. Then the law $Q := \frac{1}{2}(\delta_x + \delta_y)$ is ergodic (exercise).

Ergodic processes are important because the strong law of large numbers holds for them. More precisely:

Theorem 7.3 (Birkhoff's Ergodic Theorem for Ergodic Processes)**.** *Let Q be an ergodic law on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$. If $f : F^{\mathbb{N}} \to \mathbb{R}$ is measurable and Q-integrable (i.e., $f \in L^1(\mathrm{d}Q)$), we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(\vartheta^i x) \,=\, \mathrm{E}_Q(f)\,, \qquad Q(\mathrm{d}x)\text{-a.s. and in } L^1(\mathrm{d}Q)\,. \tag{7.7}$$

If $f : F^{\mathbb{N}} \to \mathbb{R}$ is measurable and semi-integrable, with $\mathrm{E}_Q(f) = \pm\infty$, relation (7.7) still holds $Q(\mathrm{d}x)$-a.s..

Another key feature of ergodic laws $Q \in \mathcal{M}_1(F^{\mathbb{N}})$ is that they enjoy remarkable entropy properties. First, the lim sup in (7.6) is actually a true limit, because the sequence $\frac{1}{n}h(Q_n)$ is non-increasing.[‡] Furthermore, relation (7.6) can be upgraded to an a.s. convergence.

[†]For an example of a shift-invariant set, consider the set of sequences $x \in F^{\mathbb{N}}$ that have a given proportion of letters, i.e. such that $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{\{x_i=a\}} = g_a$ for every $a \in A$, where $\{g_a\}_{a\in F} \in [0,1]^F$ is any fixed sequence.

[‡]This is true more generally for stationary laws, cf. the discussion after [25, Lemma I.6.7].

Theorem 7.4 (Shannon-McMillan-Breiman Theorem)**.** *For any ergodic law Q on $(F^{\mathbb{N}}, \mathcal{F}^{\otimes\mathbb{N}})$*

$$\lim_{n\to\infty} \frac{1}{n} \log Q\big((X_1,\ldots,X_n)=(x_1,\ldots,x_n)\big) \;=\; -H(Q)\,, \quad \textit{for } Q\textit{-a.e. } x\in F^{\mathbb{N}}\,, \tag{7.8}$$

where $H(Q)$ is the absolute entropy rate of Q, defined by relation (7.6) *(which holds as a true non-increasing limit). Furthermore, if $H(Q)<\infty$ then* (7.8) *holds true also in $L^1(\mathrm{d}Q)$.*

In words: for Q-typical sequences $x \in F^{\mathbb{N}}$, the Q-probability of the initial segment $(x_1,\ldots,x_n)$ decays exponentially fast in n with a non-random rate, that equals $-H(Q)$. This theorem is also known as the *Entropy Ergodic Theorem*, for obvious reasons, or *Asymptotic Equipartition Property*, for reasons that we explain in Remark 7.5.

Let us finally come to the relative entropy rate $h(Q|Q')$. To have a true limit in (7.5), the assumption that both $Q, Q' \in \mathcal{M}_1^{\mathrm{erg}}(F^{\mathbb{N}})$ are ergodic is not enough: suitable factorization properties of Q' are required. For instance, if Q is ergodic and $Q' = (q')^{\otimes\mathbb{N}}$ is i.i.d., by the ergodic theorem

$$H(Q|Q') \;:=\; \lim_{n\to\infty} \frac{1}{n}\, h(Q_n|Q'_n) \;=\; -H(Q) \;-\; \mathrm{E}_Q(\log q'(X_1))\,, \tag{7.9}$$

where we stress that $-\,\mathrm{E}_Q(\log q'(X_1)) \geq H(Q) \geq 0$.

Remark 7.5. Fix $\varepsilon > 0$ and let

$$A_{n,\varepsilon} \;:=\; \left\{ x\in F^{\mathbb{N}} :\; \frac{1}{n}\,\log\, Q\big((X_1,\ldots,X_n)=(x_1,\ldots,x_n)\big) \in (-H(Q)-\varepsilon, -H(Q)+\varepsilon) \right\}$$

be the set of "Q-entropy-typical sequences". By Theorem 7.4, $Q(A_{n,\varepsilon}) \to 1$ as $n\to\infty$ and this implies easily that for large n

$$\frac{1}{n}\,\log\,|A_{n,\varepsilon}| \;\in\; \big(H(Q)-\varepsilon, H(Q)+\varepsilon\big)\,,$$

where $|\cdot|$ denotes cardinality. Therefore we have the so-called *Asymptotic Equipartition Property*: for large n, the probability Q_n on F^n is approximately concentrated on a set of $\approx e^{H(Q)n}$ trajectories, each of which has probability $\approx e^{-H(Q)n}$.

Remark 7.6. The proof of Theorem 7.4 is immediate for i.i.d. laws $Q = q^{\otimes\mathbb{N}}$, writing

$$Q\big((X_1,\ldots,X_n)=(x_1,\ldots,x_n)\big) \;=\; \exp\left(\sum_{i=1}^n \log q(x_i)\right),$$

applying the strong law of large numbers and recalling that $H(Q) = h(q) = -\,\mathrm{E}_q(\log q(X_1))$. With analogous arguments, when Q is the law of an ergodic Markov chain, Theorem 7.4 can be deduced from the ordinary ergodic theorem, i.e., from Theorem 7.3.

For general ergodic laws Q, the first general proof of Theorem 7.4 in the case of countable space F is by Chung [15] (who extended previous results of Shannon, McMillan and Breiman), under the slightly stronger assumption $h(Q_1) < \infty$.

Possibly the most general version of Theorem 7.4 is proved in [4, Theorem 1]; note that the condition "$D_n > -\infty$" required in that theorem is always satisfied in our setting, because our reference measure is the counting measure (cf. the paragraph following Theorem 1 in [4]).

Remark 7.7. Note that the L^1 convergence in Theorem 7.4 is a direct consequence of (7.8) and of the existence of the limit in (7.6), thanks to Scheffé's Lemma.

8. Random sequences of words and letters

The key viewpoint in the variational approach of [11] is to look at the environment $\omega = \{\omega_n\}_{n\in\mathbb{N}}$ as a sequence of *letters* which are cut into *words* by the walk $(S = \{S_n\}_{n\in\mathbb{N}_0}, \mathbf{P})$. To present a part of this approach in a relatively elementary way, we need to strengthen considerably our assumptions on the disorder, restricting our attention to random variables with finite support (while we keep the same assumptions on the walk). More precisely:

Assumption 8.1. *The environment $(\{\omega_n\}_{n\in\mathbb{N}}, \mathbb{P})$ is an i.i.d. sequence of real random variables, taking values in the* finite set $E \subseteq \mathbb{R}$, *with zero mean, unit variance and (of course) finite exponential moments, cf.* (2.1)*. The law (more precisely, the discrete density) of ω_1 will be denoted by ν, that is $\nu(x) := \mathrm{P}(\omega_1 = x)$ for every $x \in E$. Without loss of generality, we assume that $\nu(x) > 0$ for all $x \in E$, that is E is the support of ν.*

Our approach can actually be performed in the case of a *countable* E, at the expense of some extra assumptions (some comments will be made below). We stress that the removal of the restriction that E is countable is possible with general more advanced techniques, cf. [7], but we will not discuss this issue.

For later convenience, let us state an application of the Borel-Cantelli lemma which will be useful later (the idea is taken from the "Sandwich Lemma" of [3]): for any sequence of positive random variables $\{W_n\}_{n\in\mathbb{N}}$, all defined on the same probability space, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log W_n \;\leq\; \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{E}(W_n)\,, \qquad \text{P-a.s.}\,. \tag{8.1}$$

8.1. The annealed law Q_0.

Recall that $E \subseteq \mathbb{R}$ is the (finite) space of values of the environment variables ω_n, and we set $\mathcal{E} := \mathcal{P}(E)$. The elements of E will be called *letters*. Let us introduce the space $(E^{\mathbb{N}}, \mathcal{E}^{\otimes\mathbb{N}})$ of *sequence of letters*, equipped with the product σ-field. For ease of notation, we assume that $\{\omega_n\}_{n\in\mathbb{N}}$ is the coordinate process on $E^{\mathbb{N}}$, that is $\omega_n(x) := x_n$ for every $x \in E^{\mathbb{N}}$, so that $\mathbb{P} = \nu^{\otimes\mathbb{N}}$ is a law on $E^{\mathbb{N}}$ (we recall that ν is the law of ω_1). The left shift operator on $E^{\mathbb{N}}$ will be denoted by ϑ.

The space of (finite) *words* is $\widetilde{E} := \bigcup_{n\in\mathbb{N}} E^n$. For a word $y = (x_1, \ldots, x_n) \in \widetilde{E}$ we denote by $|y| = n \in \mathbb{N}$ the corresponding word length. The space of *sequence of words* is $(\widetilde{E}^{\mathbb{N}}, \widetilde{\mathcal{E}}^{\otimes\mathbb{N}})$. The coordinate process on this space will be denoted by $\{Y_n\}_{n\in\mathbb{N}}$, that is $Y_n(y) := y_n$ for every $y \in \widetilde{E}^{\mathbb{N}}$, and the left shift operator will be denoted by $\widetilde{\vartheta}$.

There is an important law Q_0 of $\widetilde{E}^{\mathbb{N}}$, that is, the joint law of the words of environment letters cut by the walk under the annealed law $\mathbb{P} \otimes \mathbf{P}$. More precisely, recalling that $\{\tau_n\}_{n\in\mathbb{N}_0}$ denote the epochs of the successive returns of the walk to the state zero, cf. (2.8), we define Q_0 as the joint law of the sequence $\{(\omega_{\tau_{i-1}+1}, \ldots, \omega_{\tau_i})\}_{i\in\mathbb{N}}$ under $\mathbb{P} \otimes \mathbf{P}$, i.e., when ω and τ are drawn independently, according to $\mathbb{P}$ and $\mathbf{P}$ respectively. By the strong Markov property of $(S, \mathbf{P})$ and the i.i.d. character of $(\omega, \mathbb{P})$, it follows that under Q_0 the words are i.i.d.:

$$Q_0 \;=\; (q_0)^{\otimes\mathbb{N}}\,, \qquad q_0(n; \mathrm{d}x_1, \ldots, \mathrm{d}x_n) \;=\; K(n)\,\nu(\mathrm{d}x_1)\cdots\nu(\mathrm{d}x_n)\,. \tag{8.2}$$

There is a natural *concatenation map* $\kappa : \widetilde{E}^{\mathbb{N}} \to E^{\mathbb{N}}$, which glues together a sequence of words producing a sequence of letters, deleting the words separations. Note that κ is *not* injective. The push-forward through κ of a law $Q \in \mathcal{M}_1(\widetilde{E}^{\mathbb{N}})$ on $\mathcal{M}_1(E^{\mathbb{N}})$ will be denoted by $\kappa_Q := Q \circ \kappa^{-1}$: it represents the law of the sequence of *letters* (the environment), when the sequence of *words* is distributed according to Q. Plainly, $\kappa_{Q_0} = \mathbb{P}$.

With some natural abuse of notation, we still denote by $\kappa : \widetilde{E}^n \to \widetilde{E}$ the operator that glues a finite sequence of words into one single word (i.e., a finite sequence of letter).

8.2. **The target law Q and its entropic properties.** Let us consider an ergodic law $Q \in \mathcal{M}_1^{\mathrm{erg}}(\widetilde{E}^{\mathbb{N}})$ on the space of sequences of words, such that

$$m_Q := \mathrm{E}_Q(|Y_1|) < \infty. \tag{8.3}$$

The environment law κ_Q, that is the joint law of the letter sequence $\kappa(Y) = \kappa(Y_1, Y_2, \ldots)$ under Q, is *not* stationary (i.e. ϑ-shift invariant on $E^{\mathbb{N}}$) in general. For this reason, we introduce a stationarized version of it, that we call Ψ_Q, by size-biasing and randomizing the first word length: for $A \in \mathcal{E}^{\otimes \mathbb{N}}$ we set

$$\begin{aligned}\Psi_Q(A) &:= \frac{1}{m_Q} \mathrm{E}_Q\left[|Y_1| \left(\frac{1}{|Y_1|} \sum_{i=0}^{|Y_1|-1} \mathbb{1}_A(\vartheta^i \kappa(Y))\right)\right] = \frac{1}{m_Q} \mathrm{E}_Q\left[\sum_{i=0}^{|Y_1|-1} \mathbb{1}_A(\vartheta^i \kappa(Y))\right] \\ &= \frac{1}{m_Q} \mathrm{E}_Q\left[\sum_{i=0}^{|Y_1|-1} \mathbb{1}_A(\kappa(Y)_{i+1}, \kappa(Y)_{i+2}, \ldots)\right].\end{aligned}$$

Remark 8.2. It is not difficult to show that indeed Ψ_Q is a stationary law on $E^{\mathbb{N}}$, for any law Q on $\widetilde{E}^{\mathbb{N}}$ with $m_Q < \infty$. Furthermore, the ergodicity of Q transfers to Ψ_Q, that is if $Q \in \mathcal{M}_1^{\mathrm{erg}}(\widetilde{E}^{\mathbb{N}})$, as we assume, then also $\Psi_Q \in \mathcal{M}_1^{\mathrm{erg}}(E^{\mathbb{N}})$ (cf. [6, Remark 5] and the preceding lines). In particular, the ergodic theorem (Theorem 7.3) holds for Ψ_Q.

Note that κ_Q — that is, the law of $\kappa(Y)$ under Q, cf. (8.5) — is absolutely continuous w.r.t. Ψ_Q (denoted by $\kappa_Q \ll \Psi_Q$), because

$$\Psi_Q(A) \geq \frac{1}{m_Q} \mathrm{E}_Q[\vartheta^0 \mathbb{1}_A(\kappa(Y))] = \frac{1}{m_Q} Q(\kappa(Y) \in A) = \frac{1}{m_Q} \kappa_Q(A). \tag{8.4}$$

It follows that that the ergodic theorem for Ψ_Q applies to κ_Q as well: more precisely, for every $g : E^{\mathbb{N}} \to \mathbb{R}$ in $L^1(\mathrm{d}\Psi_Q)$ we have

$$\lim_{K \to +\infty} \frac{1}{K} \sum_{i=1}^{K} g(x_i, x_{i+1}, \ldots) = \mathrm{E}_{\Psi_Q}(g), \qquad \kappa_Q(\mathrm{d}x)\text{-a.s. and in } L^1(\mathrm{d}\kappa_Q). \tag{8.5}$$

It follows from $m_Q < \infty$, recall (8.3), and from the finiteness of the state space E that

$$\begin{aligned} &H(Q) < \infty, \qquad H(\Psi_Q) < \infty, \\ &|\mathrm{E}_Q(\log K(|Y_1|))| < \infty, \qquad \left|\mathrm{E}_{\Psi_Q}(\log \nu(\omega_1))\right| < \infty, \end{aligned} \tag{8.6}$$

as we prove in Remark 8.3 below. We stress that in the case of a countable state space E the conditions in (8.6) should be imposed explicitly.[†]

Since $Q_0 = (q_0)^{\otimes \mathbb{N}}$ is i.i.d., cf. (8.2), recalling (7.9) and applying twice the ergodic theorem, once for Q and once for κ_Q, we can write

$$H(Q|Q_0) = -H(Q) - \mathrm{E}_Q(\log K(|Y_1|)) - m_Q \, \mathrm{E}_{\Psi_Q}[\log \nu(\omega_1)], \tag{8.7}$$

and with analogous arguments

$$H(\Psi_Q|\mathbb{P}) = -H(\Psi_Q) - \mathrm{E}_{\Psi_Q}[\log \nu(\omega_1)]. \tag{8.8}$$

In particular, by (8.6), $H(Q|Q_0) < \infty$ and $H(\Psi_Q|\mathbb{P}) < \infty$.

Finally, let us specialize to our setting the Shannon-McMillan-Breiman Theorem for Q:

$$\lim_{\ell \to \infty} \frac{1}{\ell} \log Q\big((Y_1, \ldots, Y_\ell) = (y_1, \ldots, y_\ell)\big) = -H(Q), \qquad Q(\mathrm{d}y)\text{-a.s. and in } L^1(\mathrm{d}Q), \tag{8.9}$$

[†]Except for the third one, that follows always from $m_Q < \infty$, cf. Remark 8.3.

and for Ψ_Q:

$$\lim_{n\to\infty} \frac{1}{n} \log \Psi_Q\big((\omega_1,\ldots,\omega_n) = (x_1,\ldots,x_n)\big) \;=\; -H(\Psi_Q)\,, \quad \Psi_Q(\mathrm{d}x)\text{-a.s. and in } L^1(\mathrm{d}\Psi_Q)\,. \tag{8.10}$$

In the sequel we are actually going to need a modified version of this relation, namely:

$$\lim_{\ell\to\infty} \frac{1}{\ell} \log Q\big(\kappa(Y_1,\ldots,Y_\ell) = \kappa(y_1,\ldots,y_\ell)\big) \;=\; -m_Q\, H(\Psi_Q)\,, \quad \text{for } Q\text{-a.e. } y \in \widetilde{E}^{\mathbb{N}}\,. \tag{8.11}$$

as it is proved in [6, eq. (26) in Lemma 3] (cf. also Remark 8.4 below).

Remark 8.3. Let us show that, in the case of finite state space $|E| < \infty$, the conditions in (8.6) are a consequence of $m_Q < \infty$. By (2.2) we can write $|\log K(|Y_1|)| \le (const.) \log |Y_1| \le (const.)|Y_1|$, hence $|\mathrm{E}_Q(\log K(|Y_1|))| \le (const.) m_Q < \infty$. The finiteness of E yields immediately $|\mathrm{E}_{\Psi_Q}(\log \nu(\omega_1))| < \infty$ (because $\nu(x) > 0$ for all $x \in E$) and analogously $H(\Psi_Q) \le h((\Psi_Q)|_1) < \infty$, because the limit in (7.6) is non-increasing for ergodic laws, by Theorem 7.4. Analogously, to check that also $H(Q) < \infty$, it suffices to show that $h(Q_1) < \infty$. To this purpose, we decompose

$$\widetilde{E} \,=\, \{y \in \widetilde{E}:\; Q(Y_1 = y) \ge e^{-C|y|}\} \,\cup\, \{y \in \widetilde{E}:\; Q(Y_1 = y) < e^{-C|y|}\}\,,$$

for $C > 0$ large enough. Since $z \mapsto \log z$ is increasing and $z \mapsto z \log z$ is decreasing for $z \in (0,\varepsilon)$, when $\varepsilon > 0$ is small enough, we have

$$\begin{aligned} h(Q_1) \;&=\; -\sum_{y\in\widetilde{E}} Q(Y_1 = y)\, \log Q(Y_1 = y) \\ &\le\; -\sum_{y\in\widetilde{E}} Q(Y_1 = y) \log(e^{-C|y|}) \;-\; \sum_{y\in\widetilde{E}} e^{-C|y|} \log(e^{-C|y|}) \\ &=\; C\,\mathrm{E}_Q(|Y_1|) \;+\; C \sum_{n\in\mathbb{N}} n e^{-Cn} |E|^n \;<\; \infty\,, \end{aligned}$$

provided we choose $C > \log |E|$.

Remark 8.4. Relation (8.11) is not evident, but it is quite plausible. Let us give a more detailed explanation.

- Since $\kappa_Q \ll \Psi_Q$, relation (8.10) holds also for κ_Q-a.e. $x \in E^{\mathbb{N}}$. Moreover, we claim that we can replace the law Ψ_Q by κ_Q in the left hand side of (8.10), thus getting

$$\lim_{n\to\infty} \frac{1}{n} \log \kappa_Q\big((\omega_1,\ldots,\omega_n) = (x_1,\ldots,x_n)\big) \;=\; -H(\Psi_Q)\,, \quad \text{for } \kappa_Q\text{-a.e. } x \in E^{\mathbb{N}}\,. \tag{8.12}$$

  The "$\le$" follows immediately from (8.10) and (8.4). Then note that the difference between the terms in the left hand sides of (8.10) and (8.12) equals

$$\frac{1}{n} \log \frac{\Psi_Q\big((\omega_1,\ldots,\omega_n) = (x_1,\ldots,x_n)\big)}{\kappa_Q\big((\omega_1,\ldots,\omega_n) = (x_1,\ldots,x_n)\big)} \;=:\; \frac{1}{n}\log(\star)\,,$$

  hence the "$\ge$" in (8.12) follows by (8.1), because $\mathrm{E}_{\kappa_Q}(\star) \le 1$.
- If we set $x = \kappa(y)$ in (8.12), we can rewrite it as

$$\lim_{n\to\infty} \frac{1}{n} \log \kappa_Q\big((\omega_1,\ldots,\omega_n) = (\kappa(y)_1,\ldots,\kappa(y)_n)\big) \;=\; -H(\Psi_Q)\,, \quad \text{for } Q\text{-a.e. } y \in \widetilde{E}^{\mathbb{N}}\,. \tag{8.13}$$

  Taking the limit along the subsequence $n = |y_1| + \ldots + |y_\ell|$ and applying the ergodic theorem, according to which $\frac{1}{\ell}(|y_1| + \ldots + |y_\ell|) \to m_Q$ for Q-a.e. y, we obtain

$$\lim_{n\to\infty} \frac{1}{\ell} \log \kappa_Q\big((\omega_1,\ldots,\omega_{|y_1|+\ldots+|y_\ell|}) = \kappa(y_1,\ldots,y_\ell)\big) \;=\; -m_Q\, H(\Psi_Q)\,, \quad \text{for } Q\text{-a.e. } y \in \widetilde{E}^{\mathbb{N}}\,.$$

- Finally, using again the fact that $|y_1|+\ldots+|y_\ell| \approx \ell m_Q$, one can show that the law of $(\omega_1,\ldots,\omega_{|y_1|+\ldots+|y_\ell|})$ under κ_Q — that is the law of $(\kappa(Y)_1,\ldots,\kappa(Y)_{|y_1|+\ldots+|y_\ell|})$ under Q — is close to the law of $(\kappa(Y)_1,\ldots,\kappa(Y)_{\ell m_Q})$ under Q, which in turn is close to the law of $\kappa(Y_1,\ldots,Y_\ell)$ under Q, because $|Y_1| + \ldots + |Y_\ell| \approx \ell m_Q$ with high Q-probability. This is precisely what is done in [6, Lemma 3]. In this way one obtains (8.11).

8.3. **An intermediate law.** Let us fix a law Q in $\mathcal{M}_1^{\mathrm{erg}}(\widetilde{E}^{\mathbb{N}})$ with $m_Q < \infty$. We define a new law $\widehat{Q}_0$ on $\widetilde{E}^{\mathbb{N}}$, by setting

$$\widehat{Q}_0\big(Y_1 = y_1, \ldots, Y_\ell = y_\ell\big) \;:=\; K(|y_1|)\cdots K(|y_\ell|)\cdot \Psi_Q\big((\omega_1,\ldots,\omega_n) = \kappa(y_1,\ldots,y_\ell)\big)\,, \tag{8.14}$$

for all $\ell \in \mathbb{N}$ and $y_1,\ldots,y_\ell \in \widetilde{E}$ (where we set $n := |y_1| + \ldots + |y_\ell|$ for short). In words, we take a Ψ_Q-typical sequence of letters and we cut it into words through the original renewal process with step distribution $K(\cdot)$. Recall that the annealed Q_0 was defined in §8.1. In a sense, the law $\widehat{Q}_0$ provides an interpolation between Q_0 (from which it takes the law $K(\cdot)^{\otimes\mathbb{N}}$ of the renewal τ) and Q (from which it takes the law $\Psi_Q(\cdot)$ of the environment ω).

We claim that

$$H(Q|\widehat{Q}_0) \;=\; -H(Q) - \mathrm{E}_Q(\log K(|Y_1|)) + m_Q H(\Psi_Q)\,. \tag{8.15}$$

Recalling the definition (7.5) and (7.6) of relative and absolute entropy rates, it suffices to show that

$$m_Q\, H(\Psi_Q) \;=\; -\lim_{\ell\to+\infty}\frac{1}{\ell}\,\mathrm{E}_Q\left(\log\Psi_Q\big((\omega_1,\ldots,\omega_n)=\xi\big)\big|^{n=|Y_1|+\ldots+|Y_\ell|}_{\xi=\kappa(Y_1,\ldots,Y_\ell)}\right). \tag{8.16}$$

By (8.4), κ_Q — the law of $\kappa(Y)$ under Q — is absolutely continuous with respect to Ψ_Q. Then relation (8.10) holds for κ_Q-a.e. $x \in E^{\mathbb{N}}$, or equivalently, writing $x = \kappa(y)$, for Q-a.e. $y \in \widetilde{E}^{\mathbb{N}}$:

$$H(\Psi_Q) \;=\; -\lim_{n\to+\infty}\frac{1}{n}\log\Psi_Q\big((\omega_1,\ldots,\omega_n) = (\kappa(y)_1,\ldots,\kappa(y)_n)\big)\,, \quad \text{for } Q\text{-a.e. } y\in\widetilde{E}^{\mathbb{N}}\,.$$

Taking the limit along the random subsequence $n = |y_1| + \ldots + |y_\ell|$, for $\ell\in\mathbb{N}$, and noting that by the ergodic theorem $\frac{1}{\ell}(|y_1|+\ldots+|y_\ell|) \to m_Q$ for Q-a.e. $y\in\widetilde{E}^{\mathbb{N}}$, we get

$$m_Q\, H(\Psi_Q) = -\lim_{\ell\to+\infty}\frac{1}{\ell}\log\Psi_Q\big((\omega_1,\ldots,\omega_{|y_1|+\ldots+|y_\ell|}) = \kappa(y_1,\ldots,y_\ell)\big), \quad \text{for } Q\text{-a.e. } y\in\widetilde{E}^{\mathbb{N}}\,.$$

If we show that this convergence takes place also in $L^1(\mathrm{d}Q)$,[†] we can take the expected value E_Q and (8.16) is proved. Let us show the uniform integrability of the sequence. Setting for $n = |y_1| + \ldots + |y_\ell|$ for short, by (8.4) and by simple inclusion of events we can write

$$\begin{aligned}\Psi_Q\big((\omega_1,\ldots,\omega_n) = \kappa(y_1,\ldots,y_\ell)\big) \;&\geq\; \frac{1}{m_Q}\,Q\big((\kappa(Y)_1,\ldots,\kappa(Y)_n) = \kappa(y_1,\ldots,y_\ell)\big)\\ &\geq\; \frac{1}{m_Q}\,Q\big((Y_1,\ldots,Y_\ell) = (y_1,\ldots,y_\ell)\big)\,,\end{aligned} \tag{8.17}$$

hence

$$\begin{aligned} 0 \;&\leq\; -\frac{1}{\ell}\log\Psi_Q\big((\omega_1,\ldots,\omega_{|y_1|+\ldots+|y_\ell|}) = \kappa(y_1,\ldots,y_\ell)\big)\\ &\leq\; \frac{1}{\ell}\log m_Q \;-\; \frac{1}{\ell}\log Q\big((Y_1,\ldots,Y_\ell) = (y_1,\ldots,y_\ell)\big)\,.\end{aligned}$$

The last expression converges in $L^1(\mathrm{d}Q)$ by (8.9), therefore the uniform integrability follows.

[†]Incidentally, note that proving convergence in $L^1(\mathrm{d}Q)$ is not only sufficient but also necessary to justify taking the expected value E_Q, by Scheffé's lemma.

9. The key variational criterion for localization

We are going to present a derivation based on a $(\mathcal{G}, \mathcal{C})$-rare stretch strategy, cf. 4.1, of a key variational criterion for localization for pinning and copolymer models, cf. [11], without using (explicitly) large deviations theory. The main ideas of the strategy are inspired by [6, 7]. Throughout the section we make the stronger assumption that the support of the law of the environment variables ω_n is a finite set $E \subseteq \mathbb{R}$, cf. Assumption 8.1.

9.1. **Setup.** Recall that $\{\tau_n\}_{n\in\mathbb{N}_0}$ denote the epochs of the successive returns of the walk $(S, \mathbf{P})$ to the state zero, cf. (2.8). We denote by $\{Y_i(\omega,\tau)\}_{i\in\mathbb{N}}$ the words of environment read by the polymer, i.e.,

$$Y_i(\omega,\tau) \,:=\, (\omega_{\tau_{i-1}+1}, \ldots, \omega_{\tau_i})\,, \tag{9.1}$$

which are $\widetilde{E}$-valued random variables. Recall the definition of pinning and copolymer partition functions in §2.2, and of the general model $\mathsf{Z}_{N,\omega,\beta,h} = \mathsf{Z}^{\psi}_{N,\omega,\beta,h}$ that comprises both of them, cf. (2.10), where $\varphi(x,y) = 1_{\{y=0\}}$ for the pinning and $\varphi(x,y) = 1_{\{x\le 0,\, y\le 0\}}$ for the copolymer. For every $\beta \ge 0$ and $h \in \mathbb{R}$ we introduce the function $\varphi_{\beta,h} : \widetilde{E} \to \mathbb{R}$ that represents the energetic contribution of an excursion of the walk (recall (2.6)):

$$\begin{aligned}\varphi_{\beta,h}(n; x_1, \ldots, x_n) \,&:=\, \log \mathbf{E}\big[e^{\sum_{i=1}^n (\beta x_i - h)\psi(S_{[n-1,n]})}\big|\tau_1 = n\big] \\ &= \begin{cases} \beta x_n - h & \text{for pinning} \\ \log\big(\frac{1}{2}\big(1 + e^{-2\beta((x_1+\ldots+x_n)+nh)}\big)\big) & \text{for copolymer}\end{cases},\end{aligned} \tag{9.2}$$

where we have already performed the change of parameters in the copolymer case.

Remark 9.1. In the copolymer case there is a little nuisance for $n = 1$ (which is only relevant when $\mathrm{T} = 1$). In this case, formula (9.2) does not hold, because an excursion of length one lies on the x-axis, and not strictly above or strictly below it with probability $\frac{1}{2}$ each. However, for notational simplicity, in the sequel we always assume that (9.2) holds. This means that we assume $\mathrm{T} \ge 2$, or that when $\mathrm{T} = 1$ we redefine the copolymer model assigning a sign, decided by fair coin tossing, also to excursions of length one.

Note that there exist positive constants $c_1, c_2 \in (0,\infty)$ such that

$$|\varphi_{\beta,h}(n; x_1, \ldots, x_n)| \,\le\, c_1\,|\beta|\,\big(|x_1| + \ldots, +|x_n|\big) \,+\, c_2\,|h|\,n\,. \tag{9.3}$$

With these definitions, we can rewrite the *constrained* general partition function integrating on the excursions signs, that by (2.6) are i.i.d. fair coin tosses: for every $N \in \mathrm{T}\mathbb{N}$

$$\mathsf{Z}^{\mathrm{c}}_{N,\omega,\beta,h} \,=\, \mathbf{E}\big[e^{\beta\sum_{n=1}^N(\omega_n - h)\psi(S_{[n-1,n]})}\,1_{\{S_N=0\}}\big] \,=\, \mathbf{E}\big[e^{\sum_{i=1}^{\gamma_N}\varphi_{\beta,h}(Y_i(\omega,\tau))}\,1_{\{N\in\tau\}}\big]\,, \tag{9.4}$$

where we set

$$\gamma_N \,:=\, \sum_{i=1}^N 1_{\{S_i=0\}} \,=\, \max\{k \le N:\ \tau_k \le N\}\,.$$

(Incidentally, note that γ_N coincides with $\mathcal{L}_N$ in the case of pinning, recall (2.19), but not in the case of copolymer.) Relation (9.4) is the starting point of our approach.

Last but not least, we fix a law $Q \in \mathcal{M}_1^{\mathrm{erg}}(\widetilde{E}^{\mathbb{N}})$ — with the meaning of a target joint law for the distribution of the words read by the polymer — such that

$$m_Q \,:=\, \mathrm{E}_Q(|Y_1|) \,<\, \infty\,. \tag{9.5}$$

Recall the definition (8.14) of the intermediate law $\widehat{Q}_0$, which provides a sort of interpolation between Q_0 and Q. Since $|E| < \infty$, condition (9.5) implies the relations in (8.6), from which it follows in particular that

$$H(Q|Q_0) \,<\, \infty\,, \qquad H(\Psi_Q|\mathbb{P}) \,<\, \infty\,, \qquad H(Q|\widehat{Q}_0) \,<\, \infty\,, \tag{9.6}$$

cf. (8.7), (8.8) and (8.15). From these equations it also follows that

$$H(Q|\widehat{Q}_0) \,=\, H(Q|Q_0) - m_Q H(\Psi_Q|\mathbb{P})\,, \tag{9.7}$$

a relation that will be useful. Also note that by (9.5) and (9.3)

$$\mathrm{E}_Q(|\varphi_{\beta,h}(Y_1)|) \,<\, \infty\,. \tag{9.8}$$

Remark 9.2. A typical choice is to take and i.i.d. law $Q = q^{\otimes\mathbb{N}}$, for some $q \in \mathcal{M}_1(\widetilde{E})$, but correlated Q are expected to play a role.

9.2. **A (not too) heuristic description of the strategy.** The idea is to apply a $(\mathcal{G}, \mathcal{C})$-rare stretch strategy, built as follows.

- We look for atypical stretches $\omega = \omega_{[1,\ell]} \in \mathcal{A}_\ell$ that *look like a Ψ_Q-typical realization.* Since the original environment sequence ω is drawn according to $\mathbb{P}$, such stretches are exponentially unlikely, with a probability rate given by

$$\frac{1}{\ell}\log \mathbb{P}\big(\mathcal{A}_\ell\big) \,\approx\, \frac{1}{\ell}\log \mathbb{P}\big((\omega_1,\ldots,\omega_\ell) \text{ looks } \Psi_Q\text{-typical}\big) \,\approx\, -H(\Psi_Q|\mathbb{P})\,. \tag{9.9}$$

- Now we estimate the partition function $\mathsf{Z}^{\mathrm{c}}_{\ell,\hat\omega,\beta,h}$ for an atypical stretch $\hat\omega = \hat\omega_{[1,\ell]} \in \mathcal{A}_\ell$ and with size $\ell \approx k\, m_Q$, where $k \in \mathbb{N}$ is large enough. We give a lower bound on the partition function — recall (9.4) — forcing the walk to come back to zero $\approx k$ times (that is $\gamma_\ell \approx k$), *choosing the location of the returns τ in such a way that* $(Y_1(\hat\omega,\tau),\ldots,Y_k(\hat\omega,\tau))$ *looks like a Q-typical realization.* This yields

$$\mathsf{Z}_{m_Q k,\hat\omega} \gtrsim \exp\big(k\,\mathrm{E}_Q(\varphi_{\beta,h}(Y_1))\big)\,\mathbf{P}\big((Y_1(\hat\omega,\tau),\ldots,Y_k(\hat\omega,\tau)) \text{ looks } Q\text{-typical}\big) \tag{9.10}$$

  (where we have neglected the probability of $\{\gamma_\ell \approx k,\, \ell\in\tau\}$). We stress that $\mathbf{P}$ is the law of τ, while $\hat\omega$ is fixed and distributed (approximately) according to Ψ_Q.

- It remains to estimate from below the rate of exponential decay of the probability in (9.10). It turns out that for Ψ_Q-typical $\hat\omega$,

$$\frac{1}{k}\log \mathbf{P}\big((Y_1(\hat\omega,\tau),\ldots,Y_k(\hat\omega,\tau)) \text{ looks } Q\text{-typical}\big) \,\approx\, -H(Q|\widehat{Q}_0)\,. \tag{9.11}$$

  We recall that $\widehat{Q}_0$ is precisely the law of $\{Y_i(\hat\omega,\tau)\}_{i\in\mathbb{N}}$ when $(\hat\omega,\tau)$ are distributed like $\Psi_Q\otimes\mathbf{P}$, cf. (8.14). It follows that, if we average over Ψ_Q *inside* the logarithm in (9.11), we get, in analogy with (9.9),

$$\begin{aligned}&\frac{1}{k}\log\,\Psi_Q\otimes\mathbf{P}\big((Y_1(\hat\omega,\tau),\ldots,Y_k(\hat\omega,\tau)) \text{ looks } Q\text{-typical}\big)\\ &\quad= \frac{1}{k}\log\,\widehat{Q}_0\big((Y_1,\ldots,Y_k) \text{ looks } Q\text{-typical}\big) \,\approx\, -H(Q|\widehat{Q}_0)\,.\end{aligned} \tag{9.12}$$

  Now observe that by (8.1) the left hand side of (9.11) is always dominated (for k large) by the left hand side of (9.12). It is therefore remarkable that the right hand sides of these relations actually coincide. Some insight will be provided in the proof, cf. §9.4.

- Finally, putting together (9.9), (9.10) and (9.11) and recalling that $\ell \approx \kappa\, m_Q$, it is clear that we have built a $(\mathcal{G},\mathcal{C})$-rare stretch strategy with gain and cost given by

$$\mathcal{G} \;\approx\; \frac{1}{m_Q}\big(\, \mathrm{E}_Q(\varphi_{\beta,h}(Y_1)) \;-\; H(Q|\widehat{Q}_0)\big)\,, \qquad \mathcal{C} \;\approx\; H(\Psi_Q|\mathbb{P})\,. \tag{9.13}$$

9.3. **The localization criterion.** We are now going to turn the above heuristic description into a real proof. We need to justify properly relations (9.9), (9.10) and (9.11) — indeed, we even need to specify the vague concept of "looking like a typical realization" of a given law. The full details are given in §9.4.

Fix $\varepsilon > 0$ and recall that we have fixed an ergodic law $Q \in \mathcal{M}_1^{\mathrm{erg}}(\widetilde{E}^{\mathbb{N}})$ satisfying $m_Q < \infty$. It turns out to be technically convenient to define the set of good stretches $\mathcal{A}_\ell$ as follows:

$$\mathcal{A}_\ell \;:=\; \bigg\{\omega = \omega_{[1,\ell]} \in \mathbb{R}^\ell :\; \frac{1}{\ell}\log \mathsf{Z}^{\mathrm{c}}_{\ell,\omega,\beta,h} \;\geq\; \mathcal{G} \;:=\; \frac{1}{m_Q}\big(\,\mathrm{E}_Q(\varphi_{\beta,h}(Y_1)) - H(Q|\widehat{Q}_0)\big) - \varepsilon\bigg\}\,,$$

so that the gain $\mathcal{G}$ matches by construction (9.13). As we will show in §9.4, the set $\mathcal{A}_\ell$ contains indeed stretches that "look like a Ψ_Q-typical realization", in a precise sense.

We are going to show that also the cost of $\mathcal{A}_\ell$ matches (9.9): for a subsequence of $\ell \in \mathrm{T}\mathbb{N}$

$$\frac{1}{\ell}\log \mathbb{P}(\mathcal{A}_\ell) \;\geq\; -\mathcal{C}\,, \qquad \text{where} \qquad \mathcal{C} \;:=\; H(\Psi_Q|\mathbb{P}) + \varepsilon\,. \tag{9.14}$$

Therefore, by (4.1), we have localization if $\mathcal{G} - (1+\alpha)\mathcal{C} > 0$, that is if

$$\frac{1}{m_Q}\big(\,\mathrm{E}_Q(\varphi_{\beta,h}(Y_1)) - H(Q|\widehat{Q}_0) - (1+\alpha) m_Q H(\Psi_Q|\mathbb{P})\big) - 2\varepsilon \;>\; 0\,.$$

Since ε is arbitrary, we can drop it from the above relation. Recalling (9.7) and optimizing over the choice of Q, we find the following key variational criterion: we have localization — $(\beta,h) \in \mathcal{L}$, that is $h < \mathsf{h}_c(\beta)$, that is $\mathsf{F}(\beta,h) > 0$ — if

$$\boxed{\sup_{\substack{Q\in\mathcal{M}_1^{\mathrm{erg}}(\widetilde{E}^{\mathbb{N}}):\\ H(Q|Q_0)<\infty\,,\; m_Q<\infty}} \Big\{\,\mathrm{E}_Q(\varphi_{\beta,h}(Y_1)) \;-\; H(Q|Q_0) \;-\; \alpha\, m_Q\, H(\Psi_Q|\mathbb{P})\Big\} > 0\,.} \tag{9.15}$$

By (9.6), the condition $H(Q|Q_0) < \infty$ in the sup is redundant in our setting, but we have included it for the good reason that (9.15), as stated, is completely general. More precisely, the following important result from [11] holds true.

Theorem 9.3. *Relation* (9.15) *is a* necessary and sufficient *condition for localization, which holds for general environment laws $\mathbb{P}$ satisfying the assumptions in §2.1 (beyond the case of finite or countable support E).*

Proof. By Theorem 1.1 (iii) and Theorem 3.1 (ii) in [11], the critical line is (the unique value of $h = \mathsf{h}_c(\beta)$) such that $S_*^{\mathrm{que}}(\beta,\mathsf{h}_c(\beta)) = 0$, where $S_*^{\mathrm{que}}(\beta,h)$ is the left hand side in (9.15) except that the sup is extended to all shift-invariant (not necessarily ergodic) laws $Q \in \mathcal{M}_1(\widetilde{E}^{\mathbb{N}})$. However, since every shift-invariant law is a mixture of ergodic laws and the term in brackets in (9.15) is an *affine* function of Q, it follows that $S_*^{\mathrm{que}}(\beta,h)$ indeed coincides with the left hand side in (9.15). Note that $\varphi_{\beta,h}$ is a non-increasing function of h, cf. (9.2), hence $S_*^{\mathrm{que}}(\beta,h)$ is a non-increasing function of h as well. Since $S_*^{\mathrm{que}}(\beta,\mathsf{h}_c(\beta)) = 0$, as stated at the beginning of the proof, condition (9.15) yields $h < \mathsf{h}_c(\beta)$. □

Remark 9.4. The condition $H(Q|Q_0) < \infty$ ensures that $\mathrm{E}_Q(|\varphi_{\beta,h}(Y_1)|) < \infty$, as it is proved in [11, Lemma A.2]. Therefore the term in brackets in (9.15) is well-posed in the general setting.

9.4. **Proof of** (9.14)**.** The essential ideas of the proof are taken from [6, Lemma 9], where the following crucial combinatorial lemma is used implicitly.

Lemma 9.5. *Let Q be a probability on some space, on which are defined two random variables W, Z. Let $\eta > 0$ and C, D subsets of values of W, Z respectively be such that*

$$Q(W \in C) > 1 - \frac{\eta^2}{2}, \qquad Q(Z \in D) > 1 - \frac{\eta}{2}.$$

Assume that the conditional probability $Q(W \in C \,|\, Z = z)$ is well defined (for instance, Z is a discrete random variable). Then there exist $D' \subseteq D$ such that

$$Q\big(Z \in D'\big) > 1 - \eta,$$

and for every $z \in D'$

$$Q\big(W \in C \,\big|\, Z = z\big) > 1 - \eta.$$

Proof. Setting $g(z) := Q(W \notin C \,|\, Z = z)$ we can write

$$\mathrm{E}_Q(g(Z)) = \int g(z)\, Q(Z \in \mathrm{d}z) = Q(W \notin C) < \frac{\eta^2}{2},$$

and by Markov's inequality

$$Q\big(g(Z) \geq \eta\big) \leq \frac{\mathrm{E}_Q(g(Z))}{\eta} \leq \frac{\eta}{2}.$$

Setting $D' := D \cap \{z : \; g(z) < \eta\}$, it follows that

$$Q\big(Z \notin D'\big) \leq Q\big(Z \notin D\big) + Q\big(g(Z) \geq \eta\big) < \frac{\eta}{2} + \frac{\eta}{2} = \eta,$$

and by construction for every $z \in D'$ we have

$$Q\big(W \notin C \,\big|\, Z = z\big) = g(z) < \eta. \qquad \square$$

We are going to apply the above lemma with

$$W = (Y_1, \ldots, Y_k), \qquad Z = \kappa(Y_1, \ldots, Y_k).$$

for $k \in \mathbb{N}$ sufficiently large (to be fixed in a moment) and for the following choice of subsets $C \subseteq \widetilde{E}^k$ and $D \subseteq \widetilde{E}$. The idea is to choose a set D of typical long words, each of which can be cut into k typical sub-words (belonging to C). We set $\varphi := \varphi_{\beta,h}$ for short.

- We choose C to be the set of $(y_1, \ldots, y_k) \in \widetilde{E}^k$ satisfying the following relations:

$$\frac{1}{k}\sum_{i=1}^{k} \varphi(y_i) \;\geq\; \mathrm{E}_Q(\varphi(Y_1)) - \frac{\varepsilon}{10}, \qquad \frac{1}{k}\sum_{i=1}^{k} \log K(|y_i|) \;\geq\; \mathrm{E}_Q(\log K(|Y_1|)) - \frac{\varepsilon}{10}, \tag{9.16}$$

$$\frac{1}{k}\log Q\big((Y_1, \ldots, Y_k) = (y_1, \ldots, y_k)\big) \;\leq\; -H(Q) + \frac{\varepsilon}{5}. \tag{9.17}$$

- We choose D to be the set of words $x = (x_1, \ldots, x_n) \in \widetilde{E}$, with $n \in \mathbb{N}$ and $x_i \in E$ (note that n is not fixed), satisfying the following relations:

$$k\big(m_Q - \varepsilon^2\big) \;<\; |x| = n \;<\; k\big(m_Q + \varepsilon^2\big), \tag{9.18}$$

$$-m_Q\, H(\Psi_Q) - \frac{\varepsilon}{5} \;\leq\; \frac{1}{k}\log Q\big(\kappa(Y_1, \ldots, Y_k) = (x_1, \ldots, x_n)\big) \;\leq\; -m_Q\, H(\Psi_Q) + \frac{\varepsilon}{5}, \tag{9.19}$$

$$\frac{1}{k}\sum_{i=1}^{n} \log \nu(x_i) \;\geq\; m_Q\, \mathrm{E}_{\Psi_Q}\big(\log \nu(x_1)\big) - \frac{\varepsilon}{5}. \tag{9.20}$$

Then for all $k \in \mathbb{N}$ large enough, say $k \geq k_0$, we can apply Lemma 9.5, with (say) $\eta = \frac{1}{2}$, i.e.

$$Q\big((Y_1, \ldots, Y_k) \in C\big) > 1 - \frac{\eta^2}{2} = \frac{7}{8}, \qquad Q\big(\kappa(Y_1, \ldots, Y_k) \in D\big) > 1 - \frac{\eta}{2} = \frac{3}{4}.$$

This is possible because by construction the above relations hold with high probability for k large: the justification for (9.16), (9.18) and (9.20) is the ergodic theorem, cf. equations (7.7) and (8.5), while the justification for (9.17) and (9.19) is the Shannon-McMillan-Breiman Theorem, cf. (8.9), and its modified version, cf. (8.11), respectively.

By Lemma 9.5, there exists $D' \subseteq D$ such that

$$Q\big(\kappa(Y_1, \ldots, Y_k) \in D'\big) > 1 - \eta = \frac{1}{2} \tag{9.21}$$

and for every $x \in D'$ we have $Q\big((Y_1, \ldots, Y_k) \in C \,\big|\, \kappa(Y_1, \ldots, Y_k) = x\big) > 1 - \eta = \frac{1}{2}$. It is convenient to write

$$D' = \bigcup_{k(m_Q - \varepsilon^2) < \ell < k(m_Q + \varepsilon^2)} \mathcal{D}_\ell, \qquad \text{where} \qquad \mathcal{D}_\ell := \{x \in D' : \ |x| = \ell\}.$$

Since there are at most $2k\varepsilon^2$ values for ℓ in the union, it follows from (9.21) that there exists $\ell = \ell_k$ such that $Q\big(\kappa(Y_1, \ldots, Y_k) \in \mathcal{D}_\ell\big) > \frac{1}{2}/(2k\varepsilon^2) = \frac{1}{4k\varepsilon^2}$. Summarizing:

- for $k \in \mathbb{N}$ sufficiently large there exist $\ell_k \leq k(m_Q + \varepsilon^2)$ and $\mathcal{D}_{\ell_k} \subseteq D \subseteq E^{\ell_k}$ such that

$$Q\big(\kappa(Y_1, \ldots, Y_k) \in \mathcal{D}_{\ell_k}\big) > \frac{1}{4k\varepsilon^2}; \tag{9.22}$$

- for every $x \in \mathcal{D}_{\ell_k} \subseteq D'$ we have

$$Q\big((Y_1, \ldots, Y_k) \in C \,\big|\, \kappa(Y_1, \ldots, Y_k) = x\big) > \frac{1}{2}. \tag{9.23}$$

We are almost done. Recall that we aim at (9.14), for which it is enough to show that

$$\liminf_{k \to \infty} \frac{1}{\ell_k} \log \mathbb{P}(\mathcal{A}_{\ell_k}) \; > \; -H(\Psi_Q|\mathbb{P}) - \varepsilon. \tag{9.24}$$

To complete the proof, we show that $\mathcal{A}_{\ell_k} \supseteq \mathcal{D}_{\ell_k}$ and that (9.24) holds for $\mathcal{D}_{\ell_k}$.

Step 1. Let us show that $\mathcal{A}_{\ell_k} \supseteq \mathcal{D}_{\ell_k}$. For any fixed $x = (x_1, \ldots, x_{\ell_k}) \in \mathcal{D}_{\ell_k}$, let

$$C_x \; := \; \big\{(y_1, \ldots, y_k) \in C : \ \kappa(y_1, \ldots, y_k) = x\big\}. \tag{9.25}$$

Note that $|C_x|$ is the number of way that the word $x = (x_1, \ldots, x_{\ell_k})$ can be cut into k subwords that belong to C. By construction, for every $x \in \mathcal{D}_{\ell_k} \subseteq D$, applying (9.19), (9.23), the definition of C_x and (9.17), we have

$$\begin{aligned} e^{-(m_Q H(\Psi_Q) + \frac{\varepsilon}{5})k} \; &\leq \; Q(\kappa(Y_1, \ldots, Y_k) = x) \; < \; 2\, Q(\kappa(Y_1, \ldots, Y_k) = x, \ (Y_1, \ldots, Y_k) \in C) \\ &= \; 2 \sum_{(y_1, \ldots, y_k) \in C_x} Q((Y_1, \ldots, Y_k) = (y_1, \ldots, y_k)) \; \leq \; 2\, |C_x|\, e^{-(H(Q) - \frac{\varepsilon}{5})k}, \end{aligned}$$

hence if $k \in \mathbb{N}$ is large enough so that $2 < e^{\frac{1}{10}k\varepsilon}$ we get

$$|C_x| \; \geq \; e^{(-m_Q H(\Psi_Q) + H(Q) - \frac{\varepsilon}{2})k}. \tag{9.26}$$

Recalling (9.4), it follows that for every $x \in \mathcal{D}_{\ell_k}$ we have the right lower bound on the partition function: in fact, recalling (9.16) and (8.15),

$$\begin{aligned}\mathsf{Z}_{\ell_k,x,\beta,h} &\ge \sum_{(y_1,\dots,y_k)\in C_x} \prod_{i=1}^{k} e^{\varphi(y_i)}\, K(|y_i|) \ge e^{(\mathrm{E}_Q(\varphi(Y_1))+\mathrm{E}_Q(\log K(|Y_1|))-\frac{\varepsilon}{5})k}\, |C_x| \\ &\ge e^{(\mathrm{E}_Q(\varphi(Y_1))+\mathrm{E}_Q(\log K(|Y_1|))+H(Q)-m_Q H(\Psi_Q)-\frac{4}{5}\varepsilon)k} = e^{(\mathrm{E}_Q(\varphi(Y_1))-H(Q|\widehat{Q}_0)-\frac{4}{5}\varepsilon)k},\end{aligned}$$

and since $\ell_k \le k(m_Q+\varepsilon^2)$ by construction,

$$\frac{1}{\ell_k}\log \mathsf{Z}_{\ell_k,x,\beta,h} \ge \frac{\mathrm{E}_Q(\varphi(Y_1)) - H(Q|\widehat{Q}_0) - \frac{4}{5}\varepsilon}{m_Q+\varepsilon^2} \ge \mathrm{E}_Q(\varphi(Y_1)) - H(Q|\widehat{Q}_0) - \varepsilon\,,$$

provided $\varepsilon > 0$ is chosen small enough. This shows that indeed $\mathcal{D}_{\ell_k} \subseteq \mathcal{A}_{\ell_k}$.

Step 2. Let us show that (9.24) holds true with $\mathcal{D}_{\ell_k}$ in place of $C_{\varepsilon,k}$. We have

$$\mathbb{P}(\mathcal{D}_{\ell_k}) = \sum_{x=(x_1,\dots,x_{\ell_k})\in\mathcal{D}_{\ell_k}} \prod_{i=1}^{\ell_k} \nu(x_i)\,,$$

hence by (9.20) and (9.19) we can write

$$\begin{aligned}\mathbb{P}(\mathcal{D}_{\ell_k}) &\ge e^{(m_Q\,\mathrm{E}_{\Psi_Q}(\log\nu(x_1))-\frac{\varepsilon}{5})k} \sum_{x=(x_1,\dots,x_{\ell_k})\in\mathcal{D}_{\ell_k}} \frac{Q(\kappa(Y_1,\dots,Y_k)=(x_1,\dots,x_{\ell_k}))}{e^{(-m_Q H(\Psi_Q)+\frac{\varepsilon}{5})k}} \\ &= e^{(m_Q H(\Psi_Q)+m_Q\,\mathrm{E}_{\Psi_Q}(\log\nu(x_1))-\frac{2}{5}\varepsilon)k}\, Q\big(\kappa(Y_1,\dots,Y_k)\in\mathcal{D}_{\ell_k}\big)\,.\end{aligned}$$

Recalling (8.8) and (9.22), it follows that

$$\mathbb{P}(\mathcal{D}_{\ell_k}) \ge e^{-(m_Q\,H(\Psi_Q|\mathbb{P})+\frac{2}{5}\varepsilon)k}\, \frac{1}{4k\varepsilon^2}\,.$$

Since $\ell_k \le k(m_Q+\varepsilon^2)$, it follows that for k large enough and ε small enough

$$\frac{1}{\ell_k}\log\mathbb{P}(\mathcal{D}_{\ell_k}) \ge \frac{1}{m_Q+\varepsilon^2}\bigg(-m_Q\,H(\Psi_Q|\mathbb{P}) - \frac{2}{5}\varepsilon - \frac{1}{k}\log(4k\varepsilon^2)\bigg) \ge -H(\Psi_Q|\mathbb{P}) - \varepsilon\,,$$

hence (9.24) holds true with $\mathcal{D}_{\ell_k}$ in place of $C_{\varepsilon,k}$. □

10. The MBG lower bound revisited

From this section on, we remove the assumption that the variables ω_n have a finite (or countable) support $E \subseteq \mathbb{R}$, cf. Assumption 8.1. In fact, by Theorem 9.3, the localization criterion (9.15) holds in the general framework of §2.1. Nevertheless, for clarity reasons, we set $E = \mathbb{R}$ and keep using the notation of the previous section, like $\widetilde{E}$, etc..

We are going to prove the MBG lower bound (3.7) for pinning models, cf. §10.2, and to improve it for copolymer models, showing that it is strict for every $\beta > 0$, cf. §10.3, and that it has a strictly smaller slope at $\beta = 0$ than the true critical curve $h_c^{\mathrm{cop}}(\cdot)$, cf. §10.4. Everything in this section comes from [11].

10.1. **A simplified localization criterion.** Recalling (9.7), the term in brackets in (9.15) can be rewritten as

$$\mathrm{E}_Q(\varphi_{\beta,h}(Y_1)) - (1+\alpha)H(Q|Q_0) + \alpha\, H(Q|\widehat{Q}_0)\,.$$

Since $H(Q|\widehat{Q}_0) \ge 0$, we can drop this term, obtaining a weaker but more tractable sufficient criterion for localization:

$$\sup_{\substack{Q\in\mathcal{M}_1^{\mathrm{erg}}(\widetilde{E}^{\mathbb{N}}):\\ H(Q|Q_0)<\infty\,,\ m_Q<\infty}} \Big\{ \mathrm{E}_Q(\varphi_{\beta,h}(Y_1)) \,-\, (1+\alpha)\, H(Q|Q_0) \Big\} > 0\,.$$

Restricting the sup to i.i.d. laws $Q = q^{\otimes\mathbb{N}}$, we obtain the even more tractable criterion

$$L \,:=\, \sup_{\substack{q\in\mathcal{M}_1(\widetilde{E})\\ h(q|q_0)<\infty\,,\, m_q<\infty}} \left\{ \int_{\widetilde{E}} \varphi_{\beta,h}\, \mathrm{d}q - (1+\alpha)h(q|q_0) \right\} \,>\, 0\,. \tag{10.1}$$

Fix any sequence $\{C_N\}_{N\in\mathbb{N}}$ of subsets of $\widetilde{E}$ such that $C_N \uparrow \widetilde{E}$ and $\sup_{y\in C_N}|y| < \infty$, for every $N \in \mathbb{N}$.† Recalling that $q_0(\cdot)$ was defined in (8.2), it follows from (9.3) and (2.1) that

$$\begin{aligned} \mathcal{N}_N \,&:=\, \sum_{n\in\mathbb{N}} \int_{(x_1,\ldots,x_n)\in E^n} e^{\frac{1}{1+\alpha}\varphi_{\beta,h}(n;x_1,\ldots,x_n)}\, q_0(n;\mathrm{d}x_1,\ldots,\mathrm{d}x_n)\, 1_{C_N}(n;x_1,\ldots,x_n) \\ &= \sum_{n\in\mathbb{N}} \mathbb{E}\big(e^{\frac{1}{1+\alpha}\varphi_{\beta,h}(n;\omega_1,\ldots,\omega_n)} 1_{C_N}(n;\omega_1,\ldots,\omega_n)\big)\, K(n) \,<\, \infty\,. \end{aligned} \tag{10.2}$$

Then we can introduce, for every $N \in \mathbb{N}$, a law $\widetilde{q}_N$ on $\widetilde{E}$, defined by

$$\widetilde{q}_N(n;\mathrm{d}x_1,\ldots,\mathrm{d}x_n) = \frac{1}{\mathcal{N}_N} e^{\frac{1}{1+\alpha}\varphi_{\beta,h}(n;x_1,\ldots,x_n)}\, q_0(n;\mathrm{d}x_1,\ldots,\mathrm{d}x_n)\, 1_{C_N}(n;x_1,\ldots,x_n)\,, \tag{10.3}$$

and by construction $m_{\widetilde{q}_N} < \infty$ and also $h(\widetilde{q}_N|q_0) = \frac{1}{1+\alpha}(\int \varphi_{\beta,h}\,\mathrm{d}\widetilde{q}_N - \log\mathcal{N}_N) < \infty$ (recall again (9.3)). Then taking $q = \widetilde{q}_N$ in the second line of (10.1) we obtain the lower bound

$$L \,\ge\, (1+\alpha)\log\mathcal{N}_N\,,$$

and letting $N \to \infty$ we arrive at

$$L \,\ge\, \mathcal{N}_{\beta,h} \,:=\, \sum_{n\in\mathbb{N}} \mathbb{E}\big(e^{\frac{1}{1+\alpha}\varphi_{\beta,h}(n;\omega_1,\ldots,\omega_n)}\big)\, K(n)\,. \tag{10.4}$$

It follows that localization is proved if we show that $\mathcal{N}_{\beta,h} > 1$. We stress that also $\mathcal{N}_{\beta,h} = +\infty$ is OK. The rest of this section is devoted to the analysis of $\mathcal{N}_{\beta,h}$.

10.2. **The MBG lower bound for the pinning model.** In the pinning case we have $\varphi_{\beta,h}(n;x_1,\ldots,x_n) = \beta x_n - h$, cf. (9.2), hence

$$\mathcal{N}_{\beta,h} \,=\, \sum_{n\in\mathbb{N}} \mathbb{E}\big(e^{\frac{1}{1+\alpha}(\beta\omega_n - h)}\big)\, K(n) \,=\, \sum_{n\in\mathbb{N}} e^{\{\log \mathrm{M}(\frac{1}{1+\alpha}\beta) - \frac{h}{1+\alpha}\}n}\, K(n)\,. \tag{10.5}$$

Since $\sum_{n\in\mathbb{N}} K(n) = 1$ by assumption, recalling the definition (3.8) of the curve $h^{\mathrm{MBG}}(\cdot)$ for the pinning model, we have

$$\mathcal{N}_{\beta,h} \,>\, 1 \qquad \text{if and only if} \qquad h \,<\, h^{\mathrm{MBG,pin}}(\beta) \,:=\, \big(\tfrac{1}{1+\alpha}\big)^{-1} \log \mathrm{M}\big(\tfrac{1}{1+\alpha}\beta\big)\,.$$

†For instance, just take $C_N = \{y \in E :\ |y| \le N\}$.

Therefore we have proved the MBG bound (3.7) for the pinning model:

$$h_c^{\mathrm{pin}}(\beta) \,\ge\, h^{\mathrm{MBG,pin}}(\beta)\,, \qquad \forall \beta \ge 0\,.$$

10.3. **The MBG lower bound for copolymer models is strict.** In the copolymer case, recalling cf. (9.2), we can write

$$\begin{aligned} \mathcal{N}_{\beta,h} \,&=\, \sum_{n\in\mathbb{N}} \mathbb{E}\Big[\Big\{\frac{1}{2}\Big(1+e^{-2\beta(\omega_1+\ldots+\omega_n)-2\beta h n}\Big)\Big\}^{\frac{1}{1+\alpha}}\Big]\,K(n) \\ &=\, \sum_{n\in\mathbb{N}} \mathbb{E}\Big[f_\alpha\big(e^{-2\frac{\beta}{1+\alpha}(\omega_1+\ldots+\omega_n)-2\frac{\beta}{1+\alpha}hn}\big)\Big]\,K(n)\,, \end{aligned} \tag{10.6}$$

where for $z \in [0,\infty)$ we set

$$f_\alpha(z) \,:=\, \Big\{\frac{1}{2}\big(1+z^{1+\alpha}\big)\Big\}^{1/(1+\alpha)}\,. \tag{10.7}$$

Note that

$$\begin{aligned} f_\alpha'(z) \,&=\, 2^{-\frac{1}{1+\alpha}}\, z^\alpha\, \big(1+z^{1+\alpha}\big)^{-\frac{\alpha}{1+\alpha}}\,, \\ f_\alpha''(z) \,&=\, -\alpha\, 2^{-\frac{1}{1+\alpha}}\, z^{2\alpha}\, \big(1+z^{1+\alpha}\big)^{-\frac{1+2\alpha}{1+\alpha}} \,+\, \alpha\, 2^{-\frac{1}{1+\alpha}}\, z^{\alpha-1}\, \big(1+z^{1+\alpha}\big)^{-\frac{\alpha}{1+\alpha}} \\ &=\, \alpha\, 2^{-\frac{1}{1+\alpha}}\, z^{\alpha-1}\, \big(1+z^{1+\alpha}\big)^{-\frac{1+2\alpha}{1+\alpha}}\,, \end{aligned}$$

hence $f_\alpha(\cdot)$ is strictly convex. Then it follows by Jensen's inequality that for every $\beta > 0$, $h \in \mathbb{R}$ and $n \in \mathbb{N}$ we have the *strict* lower bound

$$\begin{aligned} \mathbb{E}\Big[f_\alpha\big(e^{-2\frac{\beta}{1+\alpha}(\omega_1+\ldots+\omega_n)-2\frac{\beta}{1+\alpha}hn}\big)\Big] \,&>\, f_\alpha\big(\mathbb{E}\big[e^{-2\frac{\beta}{1+\alpha}(\omega_1+\ldots+\omega_n)-2\frac{\beta}{1+\alpha}hn}\big]\big) \\ &=\, f_\alpha\big(e^{(\log \mathrm{M}(\frac{-2\beta}{1+\alpha})-2\frac{\beta}{1+\alpha}h)n}\big)\,. \end{aligned} \tag{10.8}$$

Since $f_\alpha(0) = 1$ and $\sum_{n\in\mathbb{N}} K(n) = 1$, recalling the definition (5.2) of $h_c^{\mathrm{MBG,cop}}(\cdot)$, it follows that for $\beta > 0$ the point $(\beta, h_c^{\mathrm{MBG,cop}}(\beta))$ is in the localized region, because

$$\mathcal{N}_{\beta,h} \,>\, 1 \quad \text{for } h \,=\, h_c^{\mathrm{MBG,cop}}(\beta) \,=\, \big(\tfrac{2\beta}{1+\alpha}\big)^{-1} \log \mathrm{M}\big(\tfrac{-2\beta}{1+\alpha}\big)\,.$$

This means that *the MBG lower bound for the copolymer model is strict*:

$$h_c^{\mathrm{cop}}(\beta) \,>\, h_c^{\mathrm{MBG,cop}}(\beta)\,, \qquad \forall \beta \in (0,\infty)\,. \tag{10.9}$$

Remark 10.1. What if we had considered a *terminating* renewal, i.e. $K(\infty) := 1 - \sum_{n\in\mathbb{N}} K(n) \in (0,1)$? Our approach works only for non-terminating renewal, but the terminating case is recovered considering the non-terminating renewal with law $\frac{1}{1-K(\infty)}K(n)$ and introducing a depinning $\log(1-K(\infty)) < 0$ to each excursion. So it suffice to replace $\varphi_{\beta,h}$ by $\varphi_{\beta,h} + \log(1-K(\infty))$ and the condition for localization becomes $\mathcal{N}_{\beta,h} > (1-K(\infty))^{-1/(1+\alpha)} (> 1)$, where $\mathcal{N}_{\beta,h}$ corresponds to the non-terminating renewal. Now observe that by (10.6), (10.7) and (10.8) it follows easily that $\mathcal{N}_{\beta,h} = +\infty$ for $h < h^{\mathrm{MBG,cop}}(\beta)$. Therefore the MBG-line is still a lower bound on the critical line of the copolymer model also when the underlying renewal is terminating (or, equivalently, when an arbitrary depinning term is added).

10.4. **On the critical slope of copolymer models.** We now want to show that the strict inequality (10.9) can be strengthened as $\beta \downarrow 0$ to a strict inequality between the slopes of the critical curves $h_c^{\mathrm{cop}}(\cdot)$ and $h_c^{\mathrm{MBG,cop}}(\cdot)$:

$$\liminf_{\beta\downarrow 0} \frac{h_c^{\mathrm{cop}}(\beta)}{\beta} > \lim_{\beta\downarrow 0} \frac{h^{\mathrm{MBG,cop}}(\beta)}{\beta} = \frac{1}{1+\alpha}, \qquad \forall \alpha > 0\,. \tag{10.10}$$

We resume from (10.6), in which we set $\beta = (1+\alpha)\varepsilon$ and $h = B\varepsilon$, for $B \in (0,\infty)$. Note that $B = 1$ corresponds precisely to the slope $\frac{h}{\beta} = \frac{1}{1+\alpha}$ of the MBG-line. Then to prove (10.10) it suffices to show that *for every* $\alpha > 0$ *there is* $B > 1$ *such that* $\mathcal{N}_{(1+\alpha)\varepsilon, B\varepsilon} > 1$ *for every* $\varepsilon > 0$ *small enough.* For computational simplicity, we assume that $K(n) \equiv c_K/n^{1+\alpha}$, $\mathrm{T} = 1$ and ω_n are i.i.d. $N(0,1)$, but everything carries through to the general case. Since $\omega_1 + \ldots + \omega_n \sim \sqrt{n} Z$, with $Z \sim N(0,1)$, we have

$$\begin{aligned}
\mathcal{N}_{(1+\alpha)\varepsilon, B\varepsilon} &= \sum_{n\in\mathbb{N}} \mathbb{E}\left[\left\{ \frac{1}{2}\Big(1 + e^{(1+\alpha)[-2\varepsilon\sqrt{n}Z - 2B\varepsilon^2 n]}\Big)\right\}^{\frac{1}{1+\alpha}} \right] \frac{c_K}{n^{1+\alpha}} \\
&= c_K \varepsilon^{2\alpha} \sum_{t\in\varepsilon^2\mathbb{N}} \mathbb{E}\left[\left\{ \frac{1}{2}\Big(1 + e^{(1+\alpha)[-2\sqrt{t}Z - 2Bt]}\Big)\right\}^{\frac{1}{1+\alpha}} \right] \frac{\varepsilon^2}{t^{1+\alpha}} \\
&=: c_K \varepsilon^{2\alpha} \sum_{t\in\varepsilon^2\mathbb{N}} g_{\alpha,B}(t)\, \frac{\varepsilon^2}{t^{1+\alpha}}\,.
\end{aligned} \tag{10.11}$$

Since $(x+y)^a \le x^a + y^a$ for $x, y \ge 0$ and $a \in [0,1]$, for all $z > 0$ we have

$$\left\{ \frac{1}{2}\big(1 + z^{1+\alpha}\big)\right\}^{\frac{1}{1+\alpha}} \le \frac{1}{2^{1/(1+\alpha)}} + \frac{z}{2^{1/(1+\alpha)}}\,,$$

and since $\mathbb{E}[e^{-2\sqrt{t}Z - 2Bt}] = e^{2(1-B)t}$, it follows that

$$g_{\alpha,B}(t) \le \frac{1}{2^{1/(1+\alpha)}}\big(1 + e^{2(1-B)t}\big)\,. \tag{10.12}$$

Since $z \mapsto \{\frac{1}{2}(1 + z^{1+\alpha})\}^{1/(1+\alpha)}$ is a convex function, by Jensen's inequality

$$g_{\alpha,B}(t) \ge \left\{ \frac{1}{2}\Big(1 + \big(\mathbb{E}\big[e^{-2\sqrt{t}Z - 2Bt}\big]\big)^{1+\alpha}\Big)\right\}^{\frac{1}{1+\alpha}} = \left\{ \frac{1}{2}\Big(1 + e^{2(1+\alpha)(1-B)t}\Big)\right\}^{\frac{1}{1+\alpha}}, \tag{10.13}$$

and the inequality is strict for $t > 0$.

Note incidentally that, if $B < 1$ (i.e. below the MBG-line), the function $g_{\alpha,B}(t)$ grows exponentially as $t \to \infty$ and therefore $\mathcal{N}_{(1+\alpha)\varepsilon, B\varepsilon} = +\infty$ for every $\varepsilon > 0$. We now focus on the case $B \ge 1$ (i.e. on or above the MBG-line). The function $g_{\alpha,B}(t)$ is bounded and $\mathcal{N}_{(1+\alpha)\varepsilon, B\varepsilon} < \infty$ for every $\varepsilon > 0$. Let us develop $g_{\alpha,B}(t)$ as $t \downarrow 0$. Since $(1+z)^a =$

$1 + az + \frac{1}{2}a(a-1)z^2 + o(z^2)$ as $z \to 0$, a Taylor expansion yields

$$\begin{aligned} g_{\alpha,B}(t) &= \mathbb{E}\bigg[\bigg\{\frac{1}{2}\Big(2 + (1+\alpha)[-2\sqrt{t}Z - 2Bt] + \frac{1}{2}(1+\alpha)^2(-2\sqrt{t}Z)^2 + o(t)\Big)\bigg\}^{\frac{1}{1+\alpha}}\bigg] \\ &= \mathbb{E}\bigg[\Big\{1 - (1+\alpha)\sqrt{t}Z - (1+\alpha)Bt + (1+\alpha)^2 tZ^2 + o(t)\Big)\Big\}^{\frac{1}{1+\alpha}}\bigg] \\ &= \mathbb{E}\Big[1 - \sqrt{t}Z - Bt + (1+\alpha)tZ^2 + \tfrac{1}{2}(1+\alpha)\big(\tfrac{1}{1+\alpha} - 1\big)tZ^2 + o(t)\Big] \\ &= 1 + \bigg(\frac{2+\alpha}{2} - B\bigg)t + o(t)\,, \qquad \text{as } t \downarrow 0\,. \end{aligned} \tag{10.14}$$

By Riemann sum approximation, it follows from (10.11) that

$$\lim_{\varepsilon \downarrow 0} \frac{\mathcal{N}_{(1+\alpha)\varepsilon, B\varepsilon} - 1}{\varepsilon^{2\alpha}} = c_K \int_0^\infty \frac{g_{\alpha,B}(t) - 1}{t^{1+\alpha}}\,\mathrm{d}t\,. \tag{10.15}$$

Note that the this integral is always finite in (η, ∞), for every $\eta > 0$, because $g_{\alpha,B}(\cdot)$ is bounded for $B \ge 1$. Since $g_{\alpha,B}(t) - 1 = (\frac{2+\alpha}{2} - B)t + o(t)$ as $t \downarrow 0$, by (10.14), the integral equals $+\infty$ if $B < \frac{2+\alpha}{2}$ and $\alpha > 1$. This means that, if $\alpha \ge 1$ and $\mathrm{m} < \frac{2+\alpha}{2+2\alpha}$, the copolymer with $h = \mathrm{m}\beta$ is localized for $\beta > 0$ small enough, or equivalently

$$\liminf_{\beta \downarrow 0} \frac{h_c^{\mathrm{cop}}(\beta)}{\beta} \ge \frac{2+\alpha}{2+2\alpha} > \frac{1}{1+\alpha}\,, \qquad \forall \alpha \ge 1\,.$$

It remains to consider the case $\alpha \in (0,1)$. Note that for $B = 1$ the lower bound (10.13) gives $g_{\alpha,B=1}(t) > 1$ for every $t > 0$, hence the right hand side of (10.15) is strictly positive for $B = 1$. By continuity and monotonicity, there exists a unique $\bar{B}_\alpha > 1$ such that the integral equals zero for $B = \bar{B}_\alpha$ and it is strictly positive for $B \in (1, \bar{B}_\alpha)$. In particular, if $\alpha \in (0,1)$, there exists $\bar{B}_\alpha > 1$ such that for $\mathrm{m} < \frac{\bar{B}_\alpha}{1+\alpha}$, the copolymer with $h = \mathrm{m}\beta$ is localized for $\beta > 0$ small enough, or equivalently

$$\liminf_{\beta \downarrow 0} \frac{h_c^{\mathrm{cop}}(\beta)}{\beta} \ge \frac{B_\alpha}{1+\alpha} > \frac{1}{1+\alpha}\,, \qquad \forall \alpha \in (0,1)\,.$$

This completes the proof of (10.10).

11. Disorder irrelevance for pinning models

As a further application of the rare stretch strategy in §4.1, we are going to prove that for every $\alpha \in (0, \frac{1}{2})$ there exists $\beta_0 > 0$ such that for all $\beta \in [0, \beta_0)$ we have

$$h_c^{\mathrm{pin}}(\beta) = h_c^{\mathrm{ANN,pin}}(\beta) = \log \mathrm{M}(\beta)\,, \tag{11.1}$$

following [14]. This is equivalent to showing that $(\beta, \log \mathrm{M}(\beta) - \varepsilon) \in \mathcal{L}$ for every $\varepsilon > 0$. We are going to apply the localization criterion (9.15), that we rewrite for convenience:

$$\sup_{\substack{Q \in \mathcal{M}_1^{\mathrm{erg}}(\widetilde{E}^{\mathbb{N}}):\\ H(Q|Q_0) < \infty\,,\ m_Q < \infty}} \Big\{ \mathrm{E}_Q(\varphi_{\beta,h}(Y_1)) - H(Q|Q_0) - \alpha\, m_Q\, H(\Psi_Q|\mathbb{P}) \Big\} > 0\,. \tag{11.2}$$

Recall that $\varphi_{\beta,h}(n;x_1,\ldots,x_n) = \beta x_n - h$ in the pinning case, cf. (9.2). Let us introduce the family of i.i.d. laws $\widetilde{Q}_N = (\widetilde{q}_N)^{\otimes\mathbb{N}}$, for $N \in \mathrm{T}\mathbb{N}$, with $\widetilde{q}_N \in \mathcal{M}_1(\widetilde{\mathbb{R}})$ defined by

$$\widetilde{q}_N(n;\mathrm{d}x_1,\ldots,\mathrm{d}x_n) \;=\; \frac{e^{\beta x_n}}{M(\beta)}\,\widetilde{K}_N(n)\nu(\mathrm{d}x_1)\cdots\nu(\mathrm{d}x_n)\,,$$

where we set

$$\widetilde{K}_N(n) \;:=\; K(n)\,\mathbf{1}_{\{n\le N-1\}} \;+\; K_{\ge}(N)\mathbf{1}_{\{n=N\}}\,, \qquad \text{where} \qquad K_{\ge}(N) \;:=\; \sum_{\ell\ge N} K(\ell)\,.$$

Note that $m_{\widetilde{Q}_N} \le N < \infty$. Recalling that $q_0(n;\mathrm{d}x_1,\ldots,\mathrm{d}x_n) = K(n)\nu(\mathrm{d}x_1)\cdots\nu(\mathrm{d}x_n)$, we can write

$$\begin{aligned}\mathrm{E}_{\widetilde{Q}_N}(\varphi_{\beta,h}(Y_1)) \;-\; H(\widetilde{Q}_N|Q_0) \;&=\; \mathrm{E}_{\widetilde{q}_N}(\beta x_n - h) \;-\; \mathrm{E}_{\widetilde{q}_N}\left(\log\frac{\mathrm{d}\widetilde{q}_N}{\mathrm{d}q_0}\right)\\ &=\; \log \mathrm{M}(\beta) - h + K_{\ge}(N)\log\frac{K_{\ge}(N)}{K(N)} \;\ge\; \log\mathrm{M}(\beta) - h\,.\end{aligned}$$

Choosing $h = \log \mathrm{M}(\beta) - \varepsilon$, with $\varepsilon > 0$ small but fixed, the term in brackets in (11.2) is larger than

$$\varepsilon \;-\; m_{\widetilde{Q}_N}\, H(\Psi_{\widetilde{Q}_N}|\mathbb{P})\,,$$

hence it suffices to show that

$$\lim_{N\to\infty}\Big(m_{\widetilde{Q}_N}\, H(\Psi_{\widetilde{Q}_N}|\mathbb{P})\Big) \;=\; 0\,. \tag{11.3}$$

Note that for every $\varepsilon > 0$ there exists $C > 0$ such that $K(n) \le C/n^{1+\alpha-\varepsilon}$ for every $n \in \mathbb{N}$. Then a Riemann sum approximation yields

$$m_{\widetilde{Q}_N} \;=\; \sum_{n=1}^{N-1} nK(n) \;+\; NK_{\ge}(N) \;\le\; C\sum_{n=1}^{N-1}\frac{1}{n^{\alpha-\varepsilon}} \;+\; o(1) \;\le\; C'\,N^{1-\alpha+\varepsilon}\,. \tag{11.4}$$

It remains to estimate the relative entropy term $H(\Psi_{\widetilde{Q}_N}|\mathbb{P})$. Instead of $\Psi_{\widetilde{Q}_N}$ we are going to work with $\kappa_{\widetilde{Q}_N}$ — the law of the environment sequence $\kappa(Y)$ under $\widetilde{Q}_N$ — which despite being non-stationary is easier to handle. We claim that $H(\Psi_{\widetilde{Q}_N}|\mathbb{P}) = H(\kappa_{\widetilde{Q}_N}|\mathbb{P})$, that is

$$H(\Psi_{\widetilde{Q}_N}|\mathbb{P}) \;=\; \limsup_{n\to\infty}\frac{1}{n}\,\mathrm{E}_{\kappa_{\widetilde{Q}_N}}\left[\log\frac{\mathrm{d}\kappa_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n}\right] \tag{11.5}$$

(actually we will need only the "$\le$").

To justify (11.5), one can prove that $\kappa_{\widetilde{Q}_N}$ is *Asymptotic Mean Stationary with stationary mean* $\Psi_{\widetilde{Q}_N}$, as in [14, Lemma 5.1], and then one can apply directly Theorem 4 in [3] (which also ensures that the lim sup in (11.5) is actually a true limit). An alternative, somewhat more direct proof of the "$\le$" in (11.5) is as follows. Since the law $\Psi_{\widetilde{Q}_N}$ is ergodic and $\mathbb{P}$ is an i.i.d. law, one can upgrade (7.5) to a true a.s. limit, cf. [4, Theorem 1], in analogy with the classical Shannon-McMillan-Breiman Theorem:

$$\lim_{n\to\infty}\frac{1}{n}\log\frac{\mathrm{d}\Psi_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n}(x_1,\ldots,x_n) \;=\; H(\Psi_{\widetilde{Q}_N}|\mathbb{P})\,, \qquad \text{for } \Psi_{\widetilde{Q}_N}\text{-a.e. } x\in\mathbb{R}^{\mathbb{N}}\,. \tag{11.6}$$

Since $\kappa_{\widetilde{Q}_N} \ll \Psi_{\widetilde{Q}_N}$ by (8.4), relation (11.6) holds also for $\kappa_{\widetilde{Q}_N}$-a.e. $x \in \mathbb{R}^{\mathbb{N}}$:

$$\lim_{n\to\infty}\frac{1}{n}\log\frac{\mathrm{d}\Psi_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n}(x_1,\ldots,x_n) \;=\; H(\Psi_{\widetilde{Q}_N}|\mathbb{P})\,, \qquad \text{for } \kappa_{\widetilde{Q}_N}\text{-a.e. } x\in\mathbb{R}^{\mathbb{N}}\,. \tag{11.7}$$

If replace $\Psi_{\widetilde{Q}_N}$ by $\kappa_{\widetilde{Q}_N}$ in the left hand side of (11.7), we get something larger, i.e.,

$$\limsup_{n\to\infty}\frac{1}{n}\log\frac{\mathrm{d}\kappa_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n}(x_1,\ldots,x_n) \;\ge\; H(\Psi_{\widetilde{Q}_N}|\mathbb{P})\,, \qquad \text{for } \kappa_{\widetilde{Q}_N}\text{-a.e. } x\in\mathbb{R}^{\mathbb{N}}\,, \tag{11.8}$$

as one checks applying (8.1) with

$$W_n \;=\; \frac{\mathrm{d}\Psi_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n} \bigg/ \frac{\mathrm{d}\kappa_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n}\,,$$

because $\mathrm{E}_{\kappa_{\widetilde{Q}_N}}(W_n) \le 1$. Finally, it follows easily from the expression (11.9) below othat the sequence in (11.8) is dominated by $\frac{1}{n}\sum_{i=1}^n |\beta\omega_i - \log \mathrm{M}(\beta)|$, hence it is uniformly integrable and the "$\le$" in (11.5) follows.

Note that

$$\frac{\mathrm{d}\kappa_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n} \;=\; \widetilde{\mathbf{E}}_N\Big(e^{\sum_{i=1}^n (\beta\omega_i - \log \mathrm{M}(\beta))\mathbb{1}_{\{i\in\tau\}}}\Big)\,, \tag{11.9}$$

where we denote by $\widetilde{\mathbf{P}}_N$ the law of a renewal processes τ with inter-arrival law given by

$$\widetilde{\mathbf{P}}_N(\tau_1 = n) \;:=\; \widetilde{K}_N(n)\,.$$

Applying Jensen's inequality in (11.5), we obtain

$$\begin{aligned} H(\Psi_{\widetilde{Q}_N}|\mathbb{P}) \;&\le\; \limsup_{n\to\infty} \frac{1}{n} \log \mathrm{E}_{\kappa_{\widetilde{Q}_N}}\left[\frac{\mathrm{d}\kappa_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n}\right] \;=\; \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left[\left(\frac{\mathrm{d}\kappa_{\widetilde{Q}_N}|_n}{\mathrm{d}\mathbb{P}|_n}\right)^2\right] \\ &=\; \limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}\left[\widetilde{\mathbf{E}}_N \otimes \widetilde{\mathbf{E}}'_N\Big(e^{\sum_{i=1}^n (\beta\omega_i - \log \mathrm{M}(\beta))(\mathbb{1}_{\{i\in\tau\}} + \mathbb{1}_{\{i\in\tau'\}})}\Big)\right] \\ &=\; \limsup_{n\to\infty} \frac{1}{n} \log\, \widetilde{\mathbf{E}}_N \otimes \widetilde{\mathbf{E}}'_N\Big(e^{\sum_{i=1}^n (\log \mathrm{M}(2\beta) - 2\log \mathrm{M}(\beta))\mathbb{1}_{\{i\in\tau\cap\tau'\}}}\Big)\,, \end{aligned} \tag{11.10}$$

where we denote by $(\tau', \widetilde{\mathbf{P}}'_N)$ and independent replica of $(\tau, \widetilde{\mathbf{P}}_N)$.

The intersection of the two renewals $\tau \cap \tau'$ under $\widetilde{\mathbf{P}}_N \otimes \widetilde{\mathbf{P}}'_N$ is still a renewal, whose inter-arrival law (which is not explicit) we denote by $\widetilde{K}^{(2)}_N(n)$. For later convenience, we denote by $K^{(2)}(n)$ the analogous quantity under the original law $\mathbf{P} \otimes \mathbf{P}$, and note that

$$\widetilde{K}^{(2)}_N(n) \;=\; K^{(2)}(n) \qquad \text{for every } n \le N-1\,. \tag{11.11}$$

Now observe that the last line in (11.10) is just the free energy of a homogeneous pinning model with inter-arrival law $\widetilde{K}^{(2)}_N(n)$. Setting $g(\beta) := \log \mathrm{M}(2\beta) - 2\log \mathrm{M}(\beta)$ for short (note that $g(\beta) = \beta^2 + o(\beta^2)$ as $\beta \downarrow 0$), it follows by (2.23) that

$$H(\Psi_{\widetilde{Q}_N}|\mathbb{P}) \;\le\; f_N(\beta)\,, \qquad \text{where} \qquad \sum_{n\in\mathbb{N}} \widetilde{K}^{(2)}_N(n)\, e^{-f_N(\beta)n} \;=\; e^{-g(\beta)}\,. \tag{11.12}$$

We are left with extracting estimates on $f_N(\beta)$ from this relation.

Recall that for any renewal set $\sigma = \{\sigma_n\}_{n\in\mathbb{N}_0}$ with a possibly defective inter-arrival law $\varrho(\cdot)$, for the total number of points $|\sigma|$ we have $\mathrm{P}(|\sigma| = k) = q(1-q)^k$ for all $k \in \mathbb{N}_0$, where $q := 1 - \sum_{n\in\mathbb{N}} \varrho(n)$, therefore

$$\mathrm{E}(|\sigma|) \;=\; \frac{1-q}{q} \;=\; \frac{\sum_{n\in\mathbb{N}} \varrho(n)}{1 - \sum_{n\in\mathbb{N}} \varrho(n)}\,, \tag{11.13}$$

where of course the right hand side is meant as ∞ if $\sum_{n\in\mathbb{N}} \varrho(n) = 1$, that is if the renewal is non-terminating. Therefore, *under the original law* we have

$$\chi \;:=\; \mathbf{E} \otimes \mathbf{E}'(|\tau \cap \tau'|) \;=\; \sum_{n\in\mathbb{N}} \mathbf{P} \otimes \mathbf{P}'(n \in \tau \cap \tau') \;=\; \sum_{n\in\mathbb{N}} \mathbf{P}(n \in \tau)^2 \begin{cases} = \infty & \text{if } \alpha > \frac{1}{2} \\ < \infty & \text{if } \alpha < \frac{1}{2} \end{cases}\,,$$

where the conclusion for $\alpha > 1$ follows by the ordinary renewal theorem, because $\mathbf{E}(\tau_1) < \infty$ and therefore

$$\lim_{n\to\infty,\, n\in\mathrm{T}\mathbb{N}} \mathbf{P}(n \in \tau) \;=\; \frac{1}{\mathbf{E}(\tau_1)} \;>\; 0\,,$$

while for $\alpha \in (0,1)$ we need to use local renewal theorems, cf. [16, Theorem B]:

$$\mathbf{P}(n \in \tau) \sim \frac{\alpha \sin(\pi\alpha)}{\pi} \frac{\mathrm{T}^2}{L(n)\, n^{1-\alpha}}, \qquad \text{as } n \to \infty\,,\ n \in \mathrm{T}\mathbb{N}\,.$$

Then by (11.13)

$$\sum_{n\in\mathbb{N}} K^{(2)}(n) \;=\; \frac{\chi}{1+\chi}\,, \tag{11.14}$$

where we recall that $\chi := \sum_{n\in\mathbb{N}} \mathbf{P}(n\in\tau)^2 = \sum_{n\in\mathbb{N}} \mathbf{P}(S_n = 0)^2$. We need the finiteness of χ, which explains the restriction to $\alpha < \frac{1}{2}$.

We can now conclude the proof. Set

$$\beta_0 := \sup\left\{\beta \ge 0 :\ e^{g(\beta)} \le \frac{1+\chi}{\chi}\right\} = \sup\left\{\beta \ge 0 :\ \frac{\mathrm{M}(2\beta)}{\mathrm{M}(\beta)^2} \le \frac{1+\chi}{\chi}\right\} > 0\,,$$

and choose $\beta \in [0,\beta_0)$, so that $\frac{\chi}{1+\chi} \le e^{-g(\beta)}$. By (11.12) and (11.11) we have

$$\begin{aligned} e^{-g(\beta)} &\le \sum_{n=1}^{N-1} \widetilde{K}^{(2)}_N(n) + e^{-f_N(\beta)N}\left(1 - \sum_{n=1}^{N-1} \widetilde{K}^{(2)}_N(n)\right) \\ &= \sum_{n=1}^{N-1} K^{(2)}(n) + e^{-f_N(\beta)N}\left(1 - \sum_{n=1}^{N-1} K^{(2)}(n)\right), \end{aligned}$$

therefore by (11.14)

$$f_N(\beta) \le -\frac{1}{N}\log\left(\frac{e^{-g(\beta)} - \sum_{n=1}^{N-1} K^{(2)}(n)}{1 - \sum_{n=1}^{N-1} K^{(2)}(n)}\right) \le -\frac{1}{N}\log\left(\frac{e^{-g(\beta)} - \frac{\chi}{1+\chi}}{1 - \frac{\chi}{1+\chi}}\right),$$

because $x \mapsto (e^{-g(\beta)} - x)/(1-x)$ is a non-increasing function for $x \in [0, e^{-g(\beta)}]$. Recalling (11.4) and (11.12), it follows that

$$\lim_{N\to\infty}\left(m_{\widetilde{Q}_N} H(\Psi_{\widetilde{Q}_N}|\mathbb{P})\right) \le \lim_{N\to\infty}\left(C\, N^{1-\alpha+\varepsilon}\cdot \frac{1}{N}\log\left(\frac{1-\frac{\chi}{1+\chi}}{e^{-g(\beta)} - \frac{\chi}{1+\chi}}\right)\right) = 0\,,$$

provided $\varepsilon > 0$ is chosen small enough. This shows that equation (11.3) holds true, and completes the proof.

References

[1] K.S. Alexander, *Excursions and local limit theorems for Bessel-like random walks*, Electron. J. Probab. **16** (2011), 1–44.

[2] K.S. Alexander and V. Sidoravicius, *Pinning of polymers and interfaces by random potentials*, Ann. Appl. Probab. **16** (2006) 636–669.

[3] P.H. Algoet, T.M. Cover, *A sandwich proof of the Shannon-McMillan-Breiman theorem*, Ann. Probab. **16** (1988), 899–909.

[4] A.R. Barron, *The strong ergodic theorem for densities: generalized Shannon-McMillan-Breiman theorem*, Ann. Probab. **13** (1985), 1292–1303.

[5] N.H. Bingham, C.H. Goldie and J.L. Teugels, *Regular variation*, Cambridge University Press, 1989.

[6] M. Birkner, *Conditional large deviations for a sequence of words*, Stochastic Process. Appl. **118** (2008), 703-729.

[7] M. Birkner, A. Greven, F. den Hollander, *Quenched LDP for words in a letter sequence*, Probab. Theory Relat. Fields **148** (2010), 403-456.

[8] T. Bodineau and G. Giacomin, *On the localization transition of random copolymers near selective interfaces*, J. Statist. Phys. **117** (2004), 17-34.

[9] T. Bodineau, G. Giacomin, H. Lacoin and F. L. Toninelli, *Copolymers at selective interfaces: new bounds on the phase diagram*, J. Statist. Phys. **132** (2008), 603–626.

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

[11] E. Bolthausen, F. den Hollander, A.A. Opoku, *A copolymer near a selective interface: variational characterization of the free energy*, preprint, arXiv:1110.1315v2.

[12] F. Caravenna and G. Giacomin, *The weak coupling limit of disordered copolymer models*, Ann. Probab. **38** (2010), 2322–2378.

[13] F. Caravenna, G. Giacomin and M. Gubinelli, *A Numerical Approach to Copolymers at Selective Interfaces*, J. Statist. Phys. **122** (2006), 799–832.

[14] D. Cheliotis, F. den Hollander, *Variational characterization of the critical curve for pinning of random polymers*, Ann. Probab. (to appear).

[15] K.L. Chung, *A note on the ergodic theorem of information theory*, Ann. Math. Statist. **32** (1961), 612–614.

[16] R.A. Doney, *One-sided Local Large Deviation and Renewal Theorems in the Case of Infinite Mean*, Probab. Theory Relat. Fields **107** (1997), 451–465.

[17] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, 3nd ed. (1968), John Wiley and Sons, New York.

[18] G. Giacomin, *Random polymer models*, Imperial College Press (2007).

[19] G. Giacomin, *Disorder and critical phenomena through basic probability models*, Springer (2011).

[20] G. Giacomin, F.L. Toninelli, *Smoothing Effect of Quenched Disorder on Polymer Depinning Transitions*, Commun. Math. Phys. **266** (2006), 1–16.

[21] G. Giacomin, F.L. Toninelli, *Smoothing of Depinning Transitions for Directed Polymers with Quenched Disorder*, Phys. Rev. Lett. **96** (2006), 070602.

[22] F. den Hollander, *Random Polymers*, École d'Été de Probabilités de Saint-Flour XXXVII–2007, Lecture Notes in Mathematics **1974**, Springer, 2009.

[23] C. Monthus, *On the Localization of Random Heteropolymers at the Interface Between Two Selective Solvents*, Eur. Phys. J. B **13** (2000), 111–130.

[24] N. Pétrélis, *Polymer pinning at an interface*, Stoc. Process. Appl. **116** (2006), 1600–1621.

[25] P.C. Shields, *The Ergodic Theory of Discrete Sample Paths*, American Mathematical Society (1991).

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

[27] F.L. Toninelli, *Disordered pinning models and copolymers: beyond annealed bounds*, Ann. Appl. Probab. **18** (2008), 1569–1587.

[28] F.L. Toninelli, *Coarse graining, fractional moments and the critical slope of random copolymers*, Electron. J. Probab. **14** (2009), 531–547.

Dipartimento di Matematica e Applicazioni, Università degli Studi di Milano-Bicocca, via Cozzi 53, 20125 Milano, Italy

E-mail address: francesco.caravenna@unimib.it