

Full-path localization of directed polymers*

Erik Bates[†]

Abstract

Certain polymer models are known to exhibit path localization in the sense that at low temperatures, the average fractional overlap of two independent samples from the Gibbs measure is bounded away from 0. Nevertheless, the question of where along the path this overlap takes place has remained unaddressed. In this article, we prove that on linear scales, overlap occurs along the entire length of the polymer. Namely, we consider time intervals of length εN , where $\varepsilon > 0$ is fixed but arbitrarily small. We then identify a constant number of distinguished trajectories such that the Gibbs measure is concentrated on paths having, with one of these distinguished paths, a fixed positive overlap simultaneously in every such interval. This result is obtained in all dimensions for a Gaussian random environment by using a recent non-local result as a key input.

Keywords: directed polymers; path localization; replica overlap; Gaussian disorder.

MSC2020 subject classifications: Primary 60K37, Secondary 60G15; 60G17; 82B44; 82D30, 82D60.

Submitted to EJP on July 6, 2020, final version accepted on May 4, 2021.

1 Introduction

In statistical physics, the phenomenon of *localization* refers to the tendency of disordered systems, especially at low temperatures, to revert to one of a small number of energetically favorable states, even as the size of the system diverges. Beginning with Anderson’s formative work [2], it has been a general goal to describe conditions (*e.g.* the presence of random impurities, random interaction strengths, or random geometry) under which localization occurs. Often, for a given model, a main challenge is to characterize free energy non-analyticities—which may be already difficult to rigorously detect—as separators between non-localized and localized phases. If this can be done, it gives rise to the task of more precisely quantifying the system’s behavior in either regime, the localized phase being more physically anomalous and thus harder to predict.

This paper focuses on such questions for *directed polymers in random environment*. Defined in the next section, this model was introduced by Huse and Henley [22] to study

*This research was supported by NSF grant DMS-1902734.

[†]University of Wisconsin–Madison, United States of America. E-mail: ewbates@wisc.edu

interfaces of the Ising model subject to random impurities, and later adopted in the mathematics literature by Imbrie and Spencer [23] as a model for polymer growth in random media. At low temperatures, directed polymers exhibit localization properties which have been a frequent object of study over the last forty years; a nearly complete survey is provided in the book of Comets [12], and related models are discussed in [19].

Most of the literature on localization has focused on the polymer's endpoint distribution, but very recently there has been progress in proving pathwise localization. For certain random environments at sufficiently low temperature, it is now known that if two polymers are sampled independently under the same environment, then with non-vanishing probability they will intersect for a non-vanishing fraction of their length. However, owing to the global nature of this property, the results to date provide little information on the local structure needed to produce this effect; for instance, *where* these intersections occur. The main purpose of this paper is to provide a first result in this direction, stated as Theorem 1.3 in Section 1.3.

Interestingly, the central input to the proof is a recent *non-local* path localization result from [7]. This plan of attack is natural from the perspective of random walks (from which polymers are defined), whose structure of i.i.d. increments frequently allows one to translate between local and global information. For directed polymers, however, there is no obvious renewal feature to function in the same way. Fortunately, we identify as a weak surrogate a multi-temperature free energy expression (4.3) that permits one to analyze isolated segments of the polymer. This technique is summarized in Section 1.4 and may be of independent interest.

1.1 The model: directed polymers in Gaussian environment

Let $\sigma = (\sigma_i)_{i \geq 0}$ denote simple random walk on \mathbb{Z}^d starting at the origin. We will write P to denote the law of σ in the space

$$\Sigma := \{ \sigma = (\sigma_i)_{i \geq 0} \in (\mathbb{Z}^d)^{\{0,1,\dots\}} : \sigma_0 = 0, \|\sigma_i - \sigma_{i-1}\| = 1 \text{ for all } i \geq 1 \}, \quad (1.1)$$

equipped with the standard cylindrical sigma-algebra. Expectation with respect to P will be denoted by $E(\cdot)$.

Next let $\mathbf{g} = (g(i, x) : i \geq 1, x \in \mathbb{Z}^d)$ be a collection of i.i.d. standard normal random variables, supported on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Expectation according to \mathbb{P} will be denoted by $\mathbb{E}(\cdot)$. The infinite collection \mathbf{g} is called the *disorder* or *random environment*, and defines a family of Hamiltonians on Σ ,

$$H_N(\sigma) := \sum_{i=1}^N g(i, \sigma_i), \quad N \geq 1.$$

At inverse temperature $\beta \geq 0$, the associated Gibbsian *polymer measure* is given by

$$\mu_{N,\beta}(d\sigma) := \frac{1}{Z_N(\beta)} e^{\beta H_N(\sigma)} P(d\sigma), \quad (1.2)$$

where $Z_N(\beta) := E(e^{\beta H_N(\sigma)})$ is the random normalization constant known as the *partition function*. As a function of N , the partition function grows exponentially with a limiting rate $p(\beta)$ called the *free energy*.

Theorem A. [10, Prop. 1.4] There exists a function $p: [0, \infty) \rightarrow \mathbb{R}$ such that

$$\lim_{N \rightarrow \infty} \frac{\mathbb{E} \log Z_N(\beta)}{N} = p(\beta) \quad \text{for all } \beta \geq 0. \quad (1.3)$$

Moreover, for any $u > 0$ and $\beta > 0$, we have

$$\mathbb{P}\left(\left|\frac{\log Z_N(\beta)}{N} - \frac{\mathbb{E} \log Z_N(\beta)}{N}\right| > u\right) \leq \exp\left(\frac{-Nu^2}{2\beta^2}\right). \tag{1.4}$$

Consequently, the following limit holds for every $\beta \geq 0$:

$$\lim_{N \rightarrow \infty} \frac{\log Z_N(\beta)}{N} = p(\beta) \quad \mathbb{P}\text{-a.s. and in } L^\alpha(\mathbb{P}) \text{ for all } \alpha \in [1, \infty). \tag{1.5}$$

Given this paper’s methods, the following observation will help avoid some technical concerns.

Remark 1.1. *A priori*, the validity of (1.5) might depend on the fact that the random variables defining $H_N(\cdot)$ also appear in $H_M(\cdot)$ for $M \geq N$. On the contrary, because of (1.4), the statement (1.5) is still true if one takes

$$H_N(\sigma) = \sum_{i=1}^N g_N(i, \sigma_i), \tag{1.6}$$

where now $\mathbf{g} = (g_N(i, x) : 1 \leq i \leq N, x \in \mathbb{Z}^d)$ is an i.i.d. collection even across N . Henceforth, we will take (1.6) as the definition of H_N . The distribution of $\mu_{N,\beta}$ does not change; only the joint law of $(\mu_{N,\beta})_{N \geq 1}$ is affected, and we will not be concerned with the latter object.

We will be interested in the relationship between $p(\beta)$ and the *overlap function*,

$$R(\sigma^1, \sigma^2) := \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{\sigma_i^1 = \sigma_i^2\}}, \quad \sigma^1, \sigma^2 \in \Sigma,$$

where the dependence of $R(\cdot, \cdot)$ on N is understood. The degree to which the model localizes can be measured by the typical size of $R(\sigma^1, \sigma^2)$ when σ^1 and σ^2 are sampled independently from $\mu_{N,\beta}$. For instance, if $\beta = 0$, then $\mu_{N,0}$ returns the simple random walk P , and classical results give the overlap’s rate of decay:

$$R(\sigma^1, \sigma^2) \asymp \begin{cases} N^{-1/2} & \text{if } d = 1, \\ N^{-1} \log N & \text{if } d = 2, \\ N^{-1} & \text{if } d \geq 3. \end{cases}$$

Considering that $R(\sigma^1, \sigma^2) \rightarrow 0$ as $N \rightarrow \infty$ in any one of these cases, it is a striking fact that when disorder is introduced at sufficiently large $\beta > 0$, this overlap remains bounded away from 0 (in various senses made precise in Section 1.2). As suggested earlier, the free energy provides an understanding of this dichotomy as a phase transition between high and low temperatures. In the following statements, the function $\beta^2/2$ appears because it is the logarithmic moment generating function of the standard normal distribution.

Theorem B. [17, Thm. 3.2] There exists a critical inverse temperature $\beta_c = \beta_c(d) \in [0, \infty)$ such that

$$\begin{aligned} 0 \leq \beta \leq \beta_c &\implies p(\beta) = \beta^2/2, \\ \beta > \beta_c &\implies p(\beta) < \beta^2/2. \end{aligned}$$

The high temperature phase $0 \leq \beta < \beta_c$ is thought to indicate a polymer measure still resembling simple random walk; a result to this effect is [17, Thm. 1.2]. On the other hand, in the low temperature phase $\beta > \beta_c$, the polymer measure is expected to be so attracted by favorable regions in the random environment that it concentrates near them. The question then is how to relate the condition $p(\beta) < \beta^2/2$ to this localization, as measured by the overlap function.

Remark 1.2. The function $\log Z_N(\cdot)$ is a logarithmic moment generating function and thus convex. It thus follows from (1.3) that $p(\cdot)$ is also convex and hence differentiable almost everywhere. It is believed (see [12, Conj. 6.1]) that there are actually no points of non-differentiability, and moreover that $p'(\beta) < \beta$ for all $\beta > \beta_c$. If this is true, then $p'(\beta) < \beta$ is equivalent to the low-temperature condition $p(\beta) < \beta^2/2$.

1.2 Background

The model we have defined makes sense if g is replaced by any family of disorder variables. The i.i.d. assumption is completely standard, and it is only out of methodological necessity that we have assumed Gaussianity. The Gaussian case happens to be one of the few for which some version of path localization has been rigorously established, but the phenomenon is anticipated in much greater generality.

The first path localization result for (1.2) appeared in [12, Thm. 6.1], although the relevant computation was already present in the work of Carmona and Hu [10, Lem. 7.1]. Adopting a Gaussian-integration-by-parts idea used in continuous models [14, 18] and earlier in the spin glass literature [1, 15, 26, 24], one can show¹ that if $p(\cdot)$ is differentiable at β , then

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\int_{\Sigma \times \Sigma} R(\sigma^1, \sigma^2) \mu_{N,\beta}^{\otimes 2}(d\sigma^1, d\sigma^2) \right] = 1 - \frac{p'(\beta)}{\beta}. \tag{1.7}$$

In particular, when $p'(\beta) < \beta$ (by Remark 1.2, this is the presumed characterization of low temperature), the average overlap between independent polymer paths has a nonzero limiting expectation. In other words, if σ^1 and σ^2 are sampled independently from $\mu_{N,\beta}$, then there is a nonzero chance that their fractional overlap $R(\sigma^1, \sigma^2)$ is at least some fixed positive number. For continuous models, analogous results can be found in [14, 18] as well as [13, Sec. 5.5].

For as elegantly simple as (1.7) is to prove, it only tells us that the previous sentence is true on an event of nonzero \mathbb{P} -probability. One should like said probability to be asymptotically equal to 1, meaning the specific realization of the disorder is irrelevant. Such was the advancement provided by Chatterjee [11], for sufficiently large β and a certain class of bounded random environments. In [7], Bates and Chatterjee proved an analogous (but less quantitative) statement in the Gaussian case, and then bootstrapped that result to the following, stronger one. The key feature is that the number J of distinguished paths has no dependence on the polymer length N , although the distinguished paths themselves, called $\sigma^1, \dots, \sigma^J$, are random and do depend on N .

Theorem C. [7, Thm. 1.6] Assume $\beta > 0$ is a point of differentiability for $p(\cdot)$ with $p'(\beta) < \beta$. Then for every $\varepsilon > 0$, there exist integers $J = J(\beta, \varepsilon)$, $N_* = N_*(\beta, \varepsilon)$ and a number $\delta = \delta(\beta, \varepsilon) > 0$ such that the following is true for all $N \geq N_*$. With \mathbb{P} -probability at least $1 - \varepsilon$, there are paths $\sigma^1, \dots, \sigma^J \in \Sigma$ satisfying

$$\mu_{N,\beta} \left(\bigcup_{j=1}^J \{ \sigma \in \Sigma : R(\sigma^j, \sigma) \geq \delta \} \right) \geq 1 - \varepsilon.$$

In this sense, $\mu_{N,\beta}$ concentrates on highly frequented paths and places no mass elsewhere. The distinguished trajectories $\sigma^1, \dots, \sigma^J$ (which are random and depend on N) might be called “favorite paths”, representing the preferred regions or “favorite corridors” of the polymer measure $\mu_{N,\beta}$.

While this brief overview has mentioned essentially all that has been proved about path localization (at least for the discrete model considered in this paper), much more

¹The identity (1.7) is verified in full, for a very general Gaussian disordered system, in [7, Cor. 3.10].

is known about localization of the endpoint distribution $\mu_{N,\beta}(\sigma_N \in \cdot)$. The state of the art goes well beyond the Gaussian case or even simple random walks (on the latter point, see [5, 4, 29] and references therein), and there is even a one-dimensional exactly solvable model [25] admitting an explicit limiting law for the endpoint distribution [16]. The reader is referred to [6] for a review of the literature.

Finally, a somewhat orthogonal direction of work considers directed polymers in heavy-tailed random environments, mostly in $d = 1$. In this setting, the degree of localization is much greater (e.g. [3, Thm. 2.1]), and so the interesting questions arise from taking $\beta = \beta_N \rightarrow 0$, where the rate of decay is determined by the index of the heavy tail [20, 27, 9, 8]. Further discussion can be found in [12, Sec. 6.4]; see also [28].

1.3 Main result

The goal of this article is to go beyond the single statistic $R(\sigma^1, \sigma^2)$. Although it serves as a natural gauge for localization, it does little to illuminate the geometry of localized polymers. For instance, if we know $R(\sigma^1, \sigma^2)$ is bounded away from zero, can we say something about the set of $i \in \{1, 2, \dots, N\}$ for which $\sigma_i^1 = \sigma_i^2$? Our main result addresses this question.

For integers $a \leq b$, let $\llbracket a, b \rrbracket$ denote the integer interval $\{a, a + 1, \dots, b\}$. Given $1 \leq a \leq b \leq N$, consider the *restricted overlap*,

$$R^{\llbracket a, b \rrbracket}(\sigma^1, \sigma^2) := \frac{1}{b - a + 1} \sum_{i=a}^b \mathbb{1}_{\{\sigma_i^1 = \sigma_i^2\}}, \quad \sigma^1, \sigma^2 \in \Sigma.$$

By examining these restricted overlaps, we will prove that the intersection set mentioned above is dense in $\llbracket 1, N \rrbracket$. Mirroring the language of Theorem C, we make this assertion precise as follows.

Theorem 1.3. *Assume $\beta > 0$ is a point of differentiability for $p(\cdot)$ with $p'(\beta) < \beta$. Then for every $\varepsilon > 0$, there exist integers $J = J(\beta, \varepsilon)$, $N_* = N_*(\beta, \varepsilon)$ and a number $\delta = \delta(\beta, \varepsilon) > 0$ such that the following is true for all $N \geq N_*$. With \mathbb{P} -probability at least $1 - \varepsilon$, there are paths $\sigma^1, \dots, \sigma^J \in \Sigma$ satisfying*

$$\mu_{N,\beta} \left(\bigcup_{j=1}^J \{ \sigma \in \Sigma : R^{\llbracket a, b \rrbracket}(\sigma^j, \sigma) \geq \delta \text{ whenever } b - a + 1 \geq \varepsilon N \} \right) \geq 1 - \varepsilon.$$

So at low temperatures and up to negligible events, a sample from the polymer measure localizes around one of a fixed number of distinguished paths, and this localization takes place along the entire length of the path; that is, in every interval of size at least εN . It is this latter part that is the contribution of the present article.

An important comment is that the statement of Theorem 1.3 concerns fixed β , which can be arbitrarily close to β_c . Prior to this result, it was only possible to make guarantees about localization away from the polymer's endpoint if β were sent to ∞ (c.f. [14, Eq. (4)], [18, Thm. 3.3.3 & 3.3.4], and [13, Sec. 9]). Indeed, because $p'(\cdot)$ is bounded (see [12, Prop. 2.1(iii)]), the identity (1.7) implies localization at a large fraction of times when β is sufficiently large, but even this leaves open the possibility of some linearly sized interval on which localization does not occur. Theorem 1.3 rules out this behavior at all low temperatures satisfying $p'(\beta) < \beta$.

A difficult and important question left open is the optimal dependence of J and δ on ε . The proof method in this paper likely leads to very poor bounds. Furthermore, can one prove localization on scales finer than linear? This would be a necessary step toward showing that, at least in $d = 1$, the polymer measure concentrates on a favorite corridor of $O(1)$ width; see [19, Sec. 12.9(6)].

1.4 Outline of proof

Consider a weaker version of Theorem 1.3 in which we replace general subintervals $\llbracket a, b \rrbracket$ by only “regular” subintervals: $(0, \frac{N}{L}]$, $(\frac{N}{L}, \frac{2N}{L}]$, \dots , $(\frac{(L-1)N}{L}, N]$, where L is some large integer. Our first observation is that Theorem 1.3 will be implied by this special case, which is stated as Theorem 2.1. Indeed, for a given $\varepsilon > 0$, we can choose L large enough that the regular subintervals are somewhat smaller than εN and thus actually contained in any interval I of size εN . In this way, positive overlap in the regular subintervals will imply positive overlap in I . This is the content of Section 2.

Another difficulty of Theorem 1.3 is that we demand the same distinguished path σ^j to be used in every subinterval $\llbracket a, b \rrbracket$ of appropriate size. The steps of the previous paragraph do not remove this requirement, and so our second reduction is to a version of Theorem 2.1 that allows the index j to depend on which regular subinterval $(\frac{(\ell-1)N}{L}, \frac{\ell N}{L}]$ is considered. This yet weaker result is stated as Theorem 3.1, and the reduction argument is given in Section 3. The rough idea is to concatenate segments of distinguished paths in order to produce a larger set but still of $O(1)$ size, so that whenever a path σ had intersected two distinct distinguished paths in consecutive subintervals, it will now intersect a single concatenated path in both subintervals. This procedure can be carried out by demanding slightly less overlap in each regular subinterval.

Having made these reductions, we are left to prove that for each regular subinterval, one can (with high probability) find a bounded number of paths such that a sample from the Gibbs measure will (with high probability) have non-vanishing overlap in the given subinterval with at least one of these paths. This statement could be easily proved if one were able to apply Theorem C within each subinterval and then take an appropriate union bound. The seeming obstruction is that the marginal of $\mu_{N,\beta}$ in a given subinterval is not a polymer measure of the same form as $\mu_{N,\beta}$. Moreover, this marginal depends on the environment at all times, not just those within the subinterval. Nevertheless, we can regard these marginals as polymer measures with respect to *random* reference measures. That is, we replace P in (1.2) by a random measure which, crucially, is determined entirely by the environment *outside* the given subinterval. Correspondingly, the Hamiltonian $H_N(\cdot)$ is replaced by a sum depending only on the environment inside said subinterval, the remaining disorder having been absorbed into the random reference measure.

In this setup, Theorem C still does not quite apply because it assumes a specific reference measure P . Fortunately, we can appeal to a more general result of [7] from which Theorem C was derived. We recall this general result as Theorem D in Section 5. The only hypothesis to check is that $\mu_{N,\beta}$ still admits a limiting free energy with respect to the random reference measures. To prove this fact, we introduce in Section 4 a “multi-temperature free energy” that, as a special case, can ignore the disorder in a given subinterval. Convergence of this generalized free energy is stated in Theorem 4.1 and proved using modifications of standard techniques. Finally, Theorem D is invoked in Section 5, where further technical issues are addressed en route to proving Theorem 3.1.

2 Reduction to regular subintervals

Given positive integers N and L , let $0 = n_0(N) \leq n_1(N) \leq \dots \leq n_L(N) = N$ be any sequence satisfying

$$n_\ell(N) - n_{\ell-1}(N) \in \left\{ \left\lfloor \frac{N}{L} \right\rfloor, \left\lceil \frac{N}{L} \right\rceil \right\} \quad \text{for all } \ell = 1, \dots, L. \quad (2.1)$$

We will think of L as fixed throughout, and then such a sequence will be chosen and fixed for each N . In other words, we partition the integer interval $\llbracket 1, N \rrbracket$ into L parts

of the form $\llbracket n_{\ell-1}(N) + 1, n_{\ell}(N) \rrbracket$, whose sizes are as close to equal as possible. For the sake of exposition, let us call these parts *regular subintervals*. The fractional overlap between $\sigma^1, \sigma^2 \in \Sigma$ in the ℓ^{th} subinterval will be denoted by

$$\begin{aligned} R^{(\ell)}(\sigma^1, \sigma^2) &:= R^{\llbracket n_{\ell-1}(N)+1, n_{\ell}(N) \rrbracket}(\sigma^1, \sigma^2) \\ &= \frac{1}{n_{\ell}(N) - n_{\ell-1}(N)} \sum_{i=n_{\ell-1}(N)+1}^{n_{\ell}(N)} \mathbb{1}_{\{\sigma_i^1 = \sigma_i^2\}}. \end{aligned} \tag{2.2}$$

When we are not varying N , we will simply write n_{ℓ} in place of $n_{\ell}(N)$.

The following special case of Theorem 1.3 will allow us to prove the general case.

Theorem 2.1. *Assume $\beta > 0$ is a point of differentiability for $p(\cdot)$ with $p'(\beta) < \beta$. Then for every $\varepsilon > 0$ and positive integer L , there exist integers $J = J(\beta, \varepsilon, L)$, $N_* = N_*(\beta, \varepsilon, L)$ and a number $\delta = \delta(\beta, \varepsilon, L) > 0$ such that the following is true for all $N \geq N_*$. With \mathbb{P} -probability at least $1 - \varepsilon$, there are paths $\sigma^1, \dots, \sigma^J \in \Sigma$ satisfying*

$$\mu_{N, \beta} \left(\bigcup_{j=1}^J \bigcap_{\ell=1}^L \{ \sigma \in \Sigma : R^{(\ell)}(\sigma^j, \sigma) \geq \delta \} \right) \geq 1 - \varepsilon.$$

Given this result, we now show that Theorem 1.3 readily follows by identifying regular subintervals lying within a given interval $\llbracket a, b \rrbracket$ of size at least εN .

Proof of Theorem 1.3. Given $\varepsilon \in (0, 1]$, let $L \geq 2$ be the integer satisfying $2/L \leq \varepsilon < 2/(L-1)$. Consider any subinterval $\llbracket a, b \rrbracket \subset \llbracket 1, N \rrbracket$ of size $b - a + 1 \geq \varepsilon N$. Our choice of L guarantees that $\llbracket n_{\ell-1}(N) + 1, n_{\ell}(N) \rrbracket \subset \llbracket a, b \rrbracket$ for some $\ell \in \llbracket 1, L \rrbracket$.

Let us first address the case when $b - a + 1 \leq 2\varepsilon N + 1$. In particular, assuming $N \geq L$, we have

$$b - a + 1 < \frac{4N}{L-1} + 1 \leq \frac{8N}{L} + 1 \leq \frac{9N}{L}.$$

Asymptotically we know $(n_{\ell} - n_{\ell-1}) \sim N/L$ as $N \rightarrow \infty$, but let us just use the trivial bounds

$$\frac{N}{2L} \leq \left\lfloor \frac{N}{L} \right\rfloor \leq n_{\ell} - n_{\ell-1} \leq \left\lceil \frac{N}{L} \right\rceil \leq \frac{2N}{L} \quad \text{for all } N \geq L. \tag{2.3}$$

For any $\sigma^1, \sigma^2 \in \Sigma$, the inclusion $\llbracket n_{\ell-1} + 1, n_{\ell} \rrbracket \subset \llbracket a, b \rrbracket$ now gives

$$\frac{1}{n_{\ell} - n_{\ell-1}} \sum_{i=n_{\ell-1}+1}^{n_{\ell}} \mathbb{1}_{\{\sigma_i^1 = \sigma_i^2\}} \leq \frac{2L}{N} \sum_{i=a}^b \mathbb{1}_{\{\sigma_i^1 = \sigma_i^2\}} \leq \frac{18}{b - a + 1} \sum_{i=a}^b \mathbb{1}_{\{\sigma_i^1 = \sigma_i^2\}}.$$

Therefore, the following implication is true:

$$R^{(\ell)}(\sigma^j, \sigma) \geq \delta \implies R^{\llbracket a, b \rrbracket}(\sigma^j, \sigma) \geq \delta/18.$$

If $b - a + 1 > 2\varepsilon N + 1$, then we can partition $\llbracket a, b \rrbracket$ into disjoint subintervals, all having sizes at least εN but no larger than $2\varepsilon N + 1$. The argument from above applies to each of these subintervals, and so

$$R^{(\ell)}(\sigma^j, \sigma) \geq \delta \quad \text{for all } \ell = 1, \dots, L \implies R^{\llbracket a, b \rrbracket}(\sigma^j, \sigma) \geq \delta/18.$$

We can now deduce Theorem 1.3 from Theorem 2.1 by replacing δ with $\delta/18$. □

3 Reduction to independent subintervals

We continue using the notation of regular subintervals introduced in Section 2, where the task of proving Theorem 1.3 was reduced to showing Theorem 2.1. In this section, we reduce Theorem 2.1 to the following, yet weaker statement.

Theorem 3.1. *Assume $\beta > 0$ is a point of differentiability for $p(\cdot)$ with $p'(\beta) < \beta$. Then for every $\varepsilon > 0$ and positive integer L , there exist integers $J = J(\beta, \varepsilon, L)$, $N_* = N_*(\beta, \varepsilon, L)$ and a number $\delta = \delta(\beta, \varepsilon, L) > 0$ such that the following is true for all $N \geq N_*$. With \mathbb{P} -probability at least $1 - \varepsilon$, there are paths $\sigma^1, \dots, \sigma^J \in \Sigma$ satisfying*

$$\mu_{N,\beta} \left(\bigcap_{\ell=1}^L \bigcup_{j=1}^J \{ \sigma \in \Sigma : R^{(\ell)}(\sigma^j, \sigma) \geq \delta \} \right) \geq 1 - \varepsilon.$$

Assuming this result, it remains to show that the same overlapping distinguished path can be taken in all L regular subintervals (i.e. exchanging the intersection and the union displayed above). We now argue that this can be done by increasing J an $O(1)$ amount and choosing δ appropriately smaller.

Proof of Theorem 2.1. Let $\varepsilon > 0$ and a positive integer L be given. Then take J , N_* , and $\delta \in (0, 1]$ as in Theorem 3.1 so that for all $N \geq N_*$, the following event occurs with \mathbb{P} -probability at least $1 - \varepsilon$. There exists a random set of paths $\mathcal{D}_1 = \{ \sigma^1, \dots, \sigma^J \} \subset \Sigma$ such that

$$\mu_{N,\beta} \left(\bigcap_{\ell=1}^L \bigcup_{\sigma' \in \mathcal{D}_1} \{ \sigma \in \Sigma : R^{(\ell)}(\sigma', \sigma) \geq \delta \} \right) \geq 1 - \varepsilon. \tag{3.1}$$

We will henceforth assume this event occurs.

Set $K := \lceil 12/\delta \rceil$. By possibly making N_* larger, we may assume N is such that

$$\lfloor N/L \rfloor \geq K.$$

We note for later that this assumption implies

$$\left\lceil \frac{\lfloor N/L \rfloor}{K} \right\rceil \stackrel{(2.3)}{\leq} \left\lceil \frac{2N/L}{K} \right\rceil \leq \frac{3N/L}{K}, \tag{3.2}$$

and also that $\delta \leq 1$ implies

$$\frac{12}{\delta} \leq K \leq \frac{12}{\delta} + 1 \leq \frac{13}{\delta}. \tag{3.3}$$

Given \mathcal{D}_1 , let us perform the following inductive procedure.

For each $\ell = 1, \dots, L$, partition the interval $\llbracket n_{\ell-1} + 1, n_\ell \rrbracket$ into K subintervals $\llbracket m_{k-1}^{(\ell)} + 1, m_k^{(\ell)} \rrbracket$, $k = 1, \dots, K$, whose sizes are as close to equal as possible. That is, we choose a sequence

$$n_{\ell-1} = m_0^{(\ell)} \leq m_1^{(\ell)} \leq \dots \leq m_K^{(\ell)} = n_\ell \tag{3.4a}$$

satisfying

$$m_k^{(\ell)} - m_{k-1}^{(\ell)} \in \left\{ \left\lfloor \frac{n_\ell - n_{\ell-1}}{K} \right\rfloor, \left\lceil \frac{n_\ell - n_{\ell-1}}{K} \right\rceil \right\} \text{ for all } k = 1, \dots, K. \tag{3.4b}$$

For $\ell \leq L - 1$, set $\mathcal{D}_{\ell+1} = \mathcal{D}_\ell \cup \mathcal{D}_\ell^+$, where the supplementary set \mathcal{D}_ℓ^+ is formed as follows. Consider every ordered pair $(\sigma', \sigma'') \in \mathcal{D}_\ell \times \mathcal{D}_\ell$ with $\sigma' \neq \sigma''$. For each $k \in \llbracket 1, K - 1 \rrbracket$, and each $t \in \llbracket n_\ell + 1, n_{\ell+1} \rrbracket$, determine whether there exists a nearest-neighbor path between

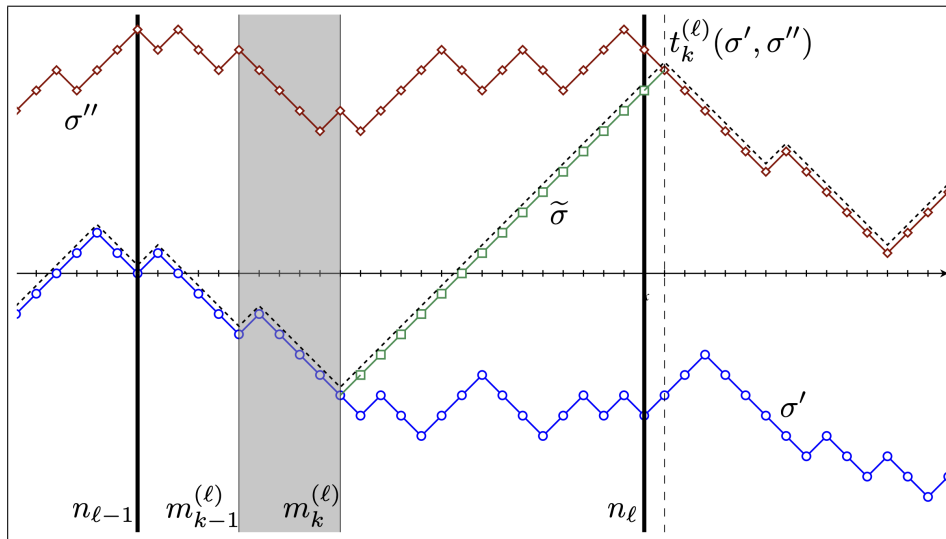


Figure 1: Example construction of $C_k^{(\ell)}(\sigma', \sigma'')$ in $d = 1$. Time is visualized in the horizontal direction. The path $\sigma' \in \mathcal{D}_\ell$ is shown as the blue solid curve with circles, $\sigma'' \in \mathcal{D}_\ell$ as the red solid curve with diamonds, and the connecting path $\tilde{\sigma}$ as the green solid curve with squares. The concatenated path $C_k^{(\ell)}(\sigma', \sigma'') \in \mathcal{D}_{\ell+1}$ is displayed as the dashed trajectory, with a dashed vertical line marking the time $t_k^{(\ell)}(\sigma', \sigma'')$ of earliest possible connection from $\sigma'_{m_k^{(\ell)}}$ to σ'' terminating in the time interval $\llbracket n_\ell + 1, n_{\ell+1} \rrbracket$.

$\sigma'_{m_k^{(\ell)}}$ and σ''_t consisting of exactly $t - m_k^{(\ell)}$ steps. Let $t_k^{(\ell)}(\sigma', \sigma'')$ be the minimal such t ; more formally,

$$t_k^{(\ell)}(\sigma', \sigma'') := \inf\{t \in \llbracket n_\ell + 1, n_{\ell+1} \rrbracket : \exists \sigma \in \Sigma, \sigma_{m_k^{(\ell)}} = \sigma'_{m_k^{(\ell)}}, \sigma_t = \sigma''_t\}. \quad (3.5)$$

If $t_k^{(\ell)}(\sigma', \sigma'') = \infty$, then we do nothing further. Otherwise, there is some nearest-neighbor path $\tilde{\sigma}$ connecting $\sigma'_{m_k^{(\ell)}}$ and $\sigma''_{t_k^{(\ell)}(\sigma', \sigma'')}$ in exactly $t_k^{(\ell)}(\sigma', \sigma'') - m_k^{(\ell)}$ steps. (If there are multiple such paths, then chose one according to some deterministic rule.) We include the following concatenated path, which we call $C_k^{(\ell)}(\sigma', \sigma'')$, as an element of \mathcal{D}_ℓ^+ (see Figure 1):

$$0 = \sigma'_0 \xrightarrow{\text{first } m_k^{(\ell)} \text{ steps of } \sigma'} \sigma'_{m_k^{(\ell)}} \xrightarrow{\text{follow } \tilde{\sigma}} \sigma''_{t_k^{(\ell)}(\sigma', \sigma'')} \xrightarrow{\text{continue on } \sigma''} \dots \quad (3.6)$$

Once this procedure has been performed for all ordered pairs $(\sigma', \sigma'') \in \mathcal{D}_\ell \times \mathcal{D}_\ell$ with $\sigma' \neq \sigma''$, the construction of the set \mathcal{D}_ℓ^+ is complete. Note that $|\mathcal{D}_\ell^+| \leq |\mathcal{D}_\ell|(|\mathcal{D}_\ell| - 1)(K - 1)$, and so $|\mathcal{D}_{\ell+1}| \leq K|\mathcal{D}_\ell|^2$, which leads to the upper bound

$$|\mathcal{D}_L| \leq K^{2^{L-1}-1} |\mathcal{D}_1|^{2^{L-1}} \stackrel{(3.3)}{\leq} \left(\frac{13}{\delta}\right)^{2^{L-1}-1} J^{2^{L-1}}.$$

In particular, $|\mathcal{D}_L|$ is bounded by a constant independent of N .

We now claim that

$$\bigcup_{\sigma' \in \mathcal{D}_L} \bigcap_{\ell=1}^L \{\sigma \in \Sigma : R^{(\ell)}(\sigma', \sigma) \geq \delta^2/104\} \supset \bigcap_{\ell=1}^L \bigcup_{\sigma' \in \mathcal{D}_1} \{\sigma \in \Sigma : R^{(\ell)}(\sigma', \sigma) \geq \delta\}. \quad (3.7)$$

In light of (3.1) and the earlier observation regarding the cardinality of \mathcal{D}_L , the containment (3.7) establishes the conclusion of Theorem 2.1 after replacing J by $(\frac{13}{\delta})^{2^{L-1}-1} J^{2^{L-1}}$ and δ by $\delta^2/104$. In order to prove (3.7), we reduce to the following claim.

Claim 3.2. *Suppose $\sigma \in \Sigma$ is such that for some $\ell \in \llbracket 1, L-1 \rrbracket$, there exist $\sigma', \sigma'' \in \mathcal{D}_\ell$ such that*

$$R^{(\ell)}(\sigma', \sigma) \geq \delta, \tag{3.8a}$$

$$R^{(\ell+1)}(\sigma'', \sigma) \geq \delta. \tag{3.8b}$$

Then there is $\sigma''' \in \mathcal{D}_{\ell+1}$ such that

$$R^{(\ell')}(\sigma''', \sigma) = R^{(\ell')}(\sigma', \sigma) \quad \text{for all } \ell' \in \llbracket 1, \ell-1 \rrbracket, \tag{3.9a}$$

$$R^{(\ell)}(\sigma''', \sigma) \geq \delta^2/104, \tag{3.9b}$$

$$R^{(\ell+1)}(\sigma''', \sigma) \geq \delta. \tag{3.9c}$$

Indeed, assume that Claim 3.2 holds. Any σ belonging to the right-hand side of (3.7) has $R^{(1)}(\sigma^1, \sigma) \geq \delta$ and $R^{(2)}(\sigma^2, \sigma) \geq \delta$ for some $\sigma^1, \sigma^2 \in \mathcal{D}_1$. So there is $\sigma^{1,2} \in \mathcal{D}_2$ such that $R^{(1)}(\sigma^{1,2}, \sigma) \geq \delta^2/104$ and $R^{(2)}(\sigma^{1,2}, \sigma) \geq \delta$. Since there is also $\sigma^3 \in \mathcal{D}_1 \subset \mathcal{D}_2$ satisfying $R^{(3)}(\sigma^3, \sigma) \geq \delta$, we can repeat the process to produce $\sigma^{1,2,3} \in \mathcal{D}_3$ satisfying $R^{(1)}(\sigma^{1,2,3}, \sigma), R^{(2)}(\sigma^{1,2,3}, \sigma) \geq \delta^2/104$ and $R^{(3)}(\sigma^{1,2,3}, \sigma) \geq \delta$. Continuing in this way, one arrives at $\sigma^{1,\dots,L} \in \mathcal{D}_L$ such that $R^{(\ell)}(\sigma^{1,\dots,L}, \sigma) \geq \delta^2/104$ for all $\ell \in \llbracket 1, L \rrbracket$. That is, σ belongs to the left-hand side of (3.7). \square

Proof of Claim 3.2. Let $\sigma \in \Sigma$ and assume (3.8) for some $\sigma', \sigma'' \in \mathcal{D}_\ell$. If $\sigma' = \sigma''$, then (3.9) holds by taking σ''' equal to $\sigma' = \sigma''$.

If instead $\sigma' \neq \sigma''$, we recall the subintervals of $\llbracket m_{k-1}^{(\ell)}+1, m_k^{(\ell)} \rrbracket, k \in \llbracket 1, K \rrbracket$, introduced in (3.4). We have

$$\sum_{k=1}^K \sum_{i=m_{k-1}^{(\ell)}+1}^{m_k^{(\ell)}} \mathbb{1}_{\{\sigma'_i=\sigma_i\}} = \sum_{i=n_{\ell-1}+1}^{n_\ell} \mathbb{1}_{\{\sigma'_i=\sigma_i\}} = (n_\ell - n_{\ell-1})R^{(\ell)}(\sigma', \sigma) \stackrel{(2.3)}{\geq} \frac{\delta N}{2L}.$$

Therefore, there must be at least two distinct values of $k \in \llbracket 1, K \rrbracket$ for which

$$\sum_{i=m_{k-1}^{(\ell)}+1}^{m_k^{(\ell)}} \mathbb{1}_{\{\sigma'_i=\sigma_i\}} \geq \frac{\delta N}{4LK}, \tag{3.10}$$

since otherwise we would have the contradictory bound

$$\sum_{k=1}^K \sum_{i=m_{k-1}^{(\ell)}+1}^{m_k^{(\ell)}} \mathbb{1}_{\{\sigma'_i=\sigma_i\}} < (K-1) \frac{\delta N}{4LK} + \left\lceil \frac{n_\ell - n_{\ell-1}}{K} \right\rceil \stackrel{(3.2)}{\leq} \frac{\delta N}{4L} + \frac{3N/L}{K} \stackrel{(3.3)}{\leq} \frac{\delta N}{2L}.$$

Let us call these two values k_1 and k_2 , where $k_1 < k_2$. We will now argue that the time $t_{k_1}^{(\ell)}(\sigma', \sigma'')$ defined in (3.5) is finite, and that the path $\sigma''' = \mathcal{C}_{k_1}^{(\ell)}(\sigma', \sigma'')$ defined in (3.6) satisfies (3.9).

Now, we know that $\sigma''_t = \sigma_t$ for some $t \in \llbracket n_\ell + 1, n_{\ell+1} \rrbracket$, simply by the fact that $R^{(\ell+1)}(\sigma'', \sigma) > 0$. We claim that any such t must satisfy $t \geq t_{k_1}^{(\ell)}(\sigma', \sigma'')$. (In particular, $t_{k_1}^{(\ell)}(\sigma', \sigma'')$ is finite.) Indeed, by (3.10) there is some $s \in \llbracket m_{k_2-1}^{(\ell)} + 1, m_{k_2}^{(\ell)} \rrbracket$ for which $\sigma'_s = \sigma_s$. Therefore, by following σ' from $\sigma'_{m_{k_1}^{(\ell)}}$ to $\sigma'_s = \sigma_s$, and then following σ from σ_s

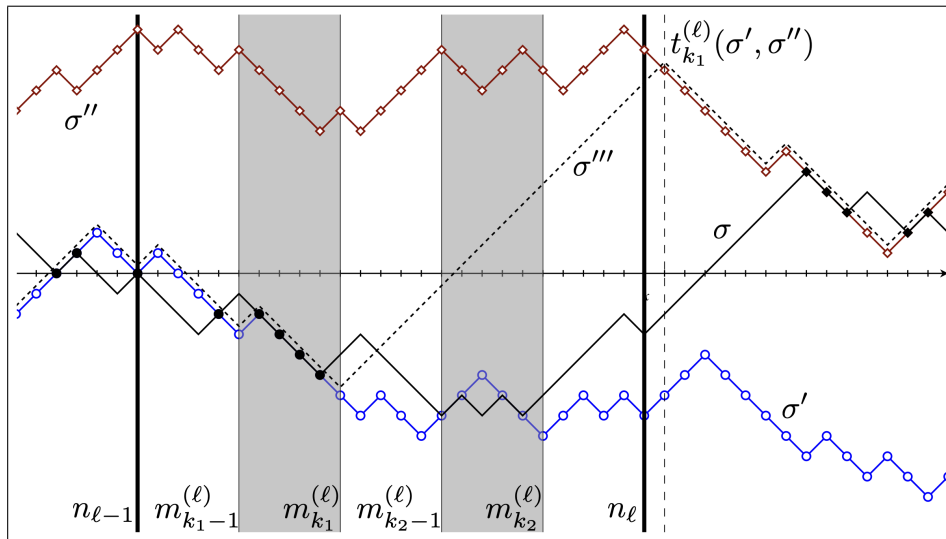


Figure 2: Overlap considerations in Claim 3.2. As in Figure 1, σ' and σ'' are displayed as solid curves with circles and diamonds, respectively, and $\sigma''' = \mathcal{C}_{k_1}^{(\ell)}(\sigma', \sigma'')$ is the dashed trajectory. The path σ under consideration is shown as a solid curve without decoration, and its intersections with σ''' are marked as filled circles or diamonds. On one hand, because σ''' agrees with σ' at least until time $m_{k_1}^{(\ell)}$, it retains all overlap with σ incurred by σ' up to that point; this observation leads to (3.9a) and (3.9b). On the other, σ is known to intersect with σ' in some later interval $\llbracket m_{k_2}^{(\ell)} - 1, m_{k_2}^{(\ell)} \rrbracket$, and so all its intersections with σ'' past time n_{ℓ} must occur after σ''' has coincided with σ'' , which occurs at time $t_{k_1}^{(\ell)}(\sigma', \sigma'')$; hence (3.9c) follows from (3.8b).

to $\sigma_t = \sigma_t''$, we will have constructed a nearest-neighbor path connecting $\sigma_{m_{k_1}^{(\ell)}}'$ to σ_t'' in exactly $t - m_{k_1}^{(\ell)}$ steps; see Figure 2. By definition (3.5), this means $t \geq t_{k_1}^{(\ell)}(\sigma', \sigma'')$.

Given that $t_{k_1}^{(\ell)}(\sigma', \sigma'')$ is finite, let σ''' be the path $\mathcal{C}_{k_1}^{(\ell)}(\sigma', \sigma'')$ constructed in (3.6). In particular, σ''' agrees with σ' up to time $m_{k_1}^{(\ell)}$, and with σ'' from time $t_{k_1}^{(\ell)}(\sigma', \sigma'')$ onward. In symbols, these are the statements

$$i \in \llbracket 1, m_{k_1}^{(\ell)} \rrbracket \implies \sigma_i''' = \sigma_i', \tag{3.11}$$

$$i \in \llbracket t_{k_1}^{(\ell)}(\sigma', \sigma''), N \rrbracket \implies \sigma_i''' = \sigma_i''. \tag{3.12}$$

Now, the argument of the previous paragraph showed that for any $t \in \llbracket n_{\ell} + 1, n_{\ell+1} \rrbracket$ such that $\sigma_t'' = \sigma_t$, we necessarily have $t \geq t_{k_1}^{(\ell)}(\sigma', \sigma'')$ and thus $\sigma_t''' = \sigma_t'' = \sigma_t$ by (3.12). In particular, we have $R^{(\ell+1)}(\sigma''', \sigma) \geq R^{(\ell+1)}(\sigma'', \sigma) \geq \delta$, thus verifying (3.9c). On the other hand, (3.9a) follows from (3.11) because $m_{k_1}^{(\ell)} > n_{\ell-1}$. Finally, to obtain (3.9b), observe that

$$\begin{aligned} R^{(\ell)}(\sigma''', \sigma) &\stackrel{(3.11)}{\geq} \frac{1}{n_{\ell} - n_{\ell-1}} \sum_{i=m_{k_1}^{(\ell)}-1+1}^{m_{k_1}^{(\ell)}} \mathbb{1}_{\{\sigma_i' = \sigma_i\}} \\ &\stackrel{(3.10)}{\geq} \frac{\delta N}{(n_{\ell} - n_{\ell-1})4LK} \stackrel{(2.3)}{\geq} \frac{\delta}{8K} \stackrel{(3.3)}{\geq} \frac{\delta^2}{104}. \end{aligned} \quad \square$$

4 Multi-temperature free energy

As outlined in Section 1.4, our proof strategy in Section 5 will require us to isolate the Hamiltonian on each regular subinterval $[[n_{\ell-1} + 1, n_\ell]]$. Mechanically, this can be done by letting the inverse temperature β depend on time and setting it equal to zero outside the interval $[[n_{\ell-1} + 1, n_\ell]]$. As it turns out, it will be easier to take the complementary route of setting the inverse temperature to zero *inside* $[[n_{\ell-1} + 1, n_\ell]]$, and keeping it unchanged outside. Either choice changes the free energy, of course, and we will need to show that a statement analogous to (1.5) still holds. It will not be any more difficult, however, to allow the inverse temperature to assume a different value on each interval $[[n_{\ell-1} + 1, n_\ell]]$, $\ell \in [[1, L]]$.

In what follows, we will use the notation $\mathcal{P}_{N,L}$ to denote the partition

$$\mathcal{P}_{N,L} : 0 = n_0(N) \leq n_1(N) \leq \dots \leq n_L(N) = N, \tag{4.1}$$

which is chosen to satisfy (2.1). Let $\beta = (\beta_1, \dots, \beta_L) \in [0, \infty)^L$ and consider the Hamiltonian

$$H_{\mathcal{P}_{N,L}}^\beta(\sigma) := \sum_{\ell=1}^L \beta_\ell \sum_{i=n_{\ell-1}+1}^{n_\ell} g_N(i, \sigma_i).$$

The associated partition function will be written as

$$Z_{\mathcal{P}_{N,L}}(\beta) := E(e^{H_{\mathcal{P}_{N,L}}^\beta(\sigma)}).$$

The concentration result stated below shows that the multi-temperature expression $\frac{1}{N} \log Z_{\mathcal{P}_{N,L}}(\beta)$ is well-approximated by an average of single-temperature free energies. Ultimately, we will need only the asymptotic statement (4.3), but with minimal extra effort we can prove (4.2) as an intermediate step.

Theorem 4.1. *For any $L \geq 1$, there exist positive constants $C_1 = C_1(L, d)$ and $c_2 = c_2(L)$ such that for any $u > 0$, $N \geq L^2$, $\beta \in [0, \infty)^L$, and $\mathcal{P}_{N,L}$ satisfying (2.1), we have*

$$\mathbb{P}\left(\left|\frac{\log Z_{\mathcal{P}_{N,L}}(\beta)}{N} - \sum_{\ell=1}^L \frac{\mathbb{E} \log Z_{n_\ell - n_{\ell-1}}(\beta_\ell)}{N}\right| > u\right) \leq C_1 N^d \exp\left(-c_2 \frac{Nu^2}{\beta_{\max}^2}\right), \tag{4.2}$$

where $\beta_{\max} := \max\{\beta_1, \dots, \beta_L, 1\}$. Consequently,

$$\lim_{N \rightarrow \infty} \frac{\log Z_{\mathcal{P}_{N,L}}(\beta)}{N} = \frac{1}{L} \sum_{\ell=1}^L p(\beta_\ell) \quad \mathbb{P}\text{-a.s. and in } L^\alpha(\mathbb{P}) \text{ for all } \alpha \in [1, \infty). \tag{4.3}$$

In the proof, we will use the following simple lemma.

Lemma 4.2. *Let X be a random variable taking values in (a, b) with mean Υ , where $-\infty \leq a < b \leq \infty$. Let $f: (a, b) \rightarrow \mathbb{R}$ be a monotone function, and denote the mean of $f(X)$ by Υ_f . If $|f(\Upsilon) - \Upsilon_f| > 2u > 0$, then the event $\{|f(X) - y| > u\}$ has positive probability for every $y \in \mathbb{R}$.*

Proof. We prove the contrapositive. That is, assume $f(X) \in [y - u, y + u]$ with probability one, for some $y \in \mathbb{R}$. Without loss of generality, we may assume f is non-decreasing; if not, we apply the argument to $-f$. If X is an almost sure constant, then $f(\Upsilon) = \Upsilon_f$. Otherwise, there exists $\delta > 0$ such that each of $\{X \leq \Upsilon - \delta\}$ and $\{X \geq \Upsilon + \delta\}$ occur with positive probability. In this case, we must have

$$y - u \leq f(\Upsilon - \delta) \leq f(\Upsilon) \leq f(\Upsilon + \delta) \leq y + u.$$

Of course, we also know $\Upsilon_f \in [y - u, y + u]$, and so $|f(\Upsilon) - \Upsilon_f| \leq 2u$. □

Proof of Theorem 4.1. We proceed by induction on L . Our inductive hypothesis is that for any $u > 0$, any integer $N \geq L^2$, and any partition

$$\mathcal{P}_{N,L} : 0 = n_0(N) < n_1(N) < \dots < n_L(N) = N$$

that is valid in the sense of (2.1), we have

$$\mathbb{P} \left(\left| \frac{\log Z_{\mathcal{P}_{N,L}}(\boldsymbol{\beta})}{N} - \sum_{\ell=1}^L \frac{\mathbb{E} \log Z_{n_\ell - n_{\ell-1}}(\beta_\ell)}{N} \right| > u \right) \leq 2L(2N+1)^d \exp \left(- \frac{Nu^2}{18L^2 \beta_{\max}^2} \right), \tag{4.4}$$

where (1.4) provides the base case of $L = 1$. Once we prove the inductive step, (4.4) will yield (4.2) with $C_1 = 3^d \cdot 2L$ and $c_2 = 1/(18L^2)$. Then, by standard arguments, it follows that

$$\lim_{N \rightarrow \infty} \left| \frac{\log Z_{\mathcal{P}_{N,L}}(\boldsymbol{\beta})}{N} - \sum_{\ell=1}^L \frac{\mathbb{E} \log Z_{n_\ell - n_{\ell-1}}(\beta_\ell)}{N} \right| = 0 \quad \mathbb{P}\text{-a.s. and in } L^\alpha(\mathbb{P}), \alpha \in [1, \infty).$$

This limit is seen to be equivalent to (4.3) once we recall (1.3). Therefore, the rest of the proof is establishing the induction. We thus assume (4.4) and consider any $\beta_{L+1} \in [0, \infty)$ and any partition

$$\mathcal{P}_{N,L+1} : 0 = n_0(N) < n_1(N) < \dots < n_L(N) < n_{L+1}(N) = N,$$

where now $N \geq (L+1)^2$.

Define the set $D_i := \{x \in \mathbb{Z}^d : P(\sigma_i = x) > 0\}$ for each integer $i \geq 1$. Observe that for any fixed realization of the disorder \mathbf{g} , if we condition on the value of σ_i for some $i \in [1, N]$, then by the Markov property of the random walk, the vectors

$$(g_N(1, \sigma_1), \dots, g_N(i, \sigma_i)) \quad \text{and} \quad (g_N(i, \sigma_i), \dots, g_N(N, \sigma_N))$$

are conditionally independent with respect to P . Using this observation when $i = n_L$, we have

$$\begin{aligned} Z_{\mathcal{P}_{N,L+1}}(\boldsymbol{\beta}, \beta_{L+1}) &= \sum_{x \in \mathbb{Z}^d} E(e^{H_{\mathcal{P}_{N,L+1}}^{(\boldsymbol{\beta}, \beta_{L+1})}(\sigma)} \mid \sigma_{n_L} = x) P(\sigma_{n_L} = x) \\ &= \sum_{x \in D_{n_L}} \left[E \left(\exp \left\{ \sum_{\ell=1}^L \beta_\ell \sum_{i=n_{\ell-1}+1}^{n_\ell} g_N(i, \sigma_i) \right\} \mid \sigma_{n_L} = x \right) P(\sigma_{n_L} = x) \right. \\ &\quad \left. \times E \left(\exp \left\{ \beta_{L+1} \sum_{i=n_L+1}^{n_{L+1}} g_N(i, \sigma_i) \right\} \mid \sigma_{n_L} = x \right) \right]. \end{aligned}$$

To condense notation, let us write

$$\begin{aligned} A(x) &:= E \left(\exp \left\{ \sum_{\ell=1}^L \beta_\ell \sum_{i=n_{\ell-1}+1}^{n_\ell} g_N(i, \sigma_i) \right\}; \sigma_{n_L} = x \right), \\ B(x) &:= E \left(\exp \left\{ \beta_{L+1} \sum_{i=n_L+1}^{n_{L+1}} g_N(i, \sigma_i) \right\} \mid \sigma_{n_L} = x \right). \end{aligned}$$

Note that

$$A := \sum_{x \in D_{n_L}} A(x) = Z_{\bar{\mathcal{P}}_{N,L}}(\beta_1, \dots, \beta_L),$$

where $\bar{\mathcal{P}}_{n_L, L}$ is the partition of $\llbracket 1, n_L \rrbracket$ into L parts induced by $\mathcal{P}_{N, L+1}$. That is,

$$\bar{\mathcal{P}}_{n_L, L} : 0 = \bar{n}_0(n_L) \leq \bar{n}_1(n_L) \leq \dots \leq \bar{n}_L(n_L) = n_L, \text{ where } \bar{n}_\ell(n_L(N)) = n_\ell(N).$$

We are thus interested in the limit of

$$\begin{aligned} \frac{\log Z_{\mathcal{P}_{N, L+1}}(\beta, \beta_{L+1})}{N} &= \frac{1}{N} \log \left(A \sum_{x \in D_{n_L}} \frac{A(x)}{A} B(x) \right) \\ &= \frac{\log Z_{\bar{\mathcal{P}}_{n_L, L}}(\beta)}{N} + \frac{1}{N} \log \sum_{x \in D_{n_L}} \frac{A(x)}{A} B(x). \end{aligned} \tag{4.5}$$

Since $N \geq (L + 1)^2$, we necessarily have $N/(L + 1) \geq L + 1$ and thus

$$\frac{\lceil N/(L + 1) \rceil}{\lfloor N/(L + 1) \rfloor} \leq \frac{L + 2}{L + 1}. \tag{4.6}$$

Therefore,

$$\begin{aligned} n_L &\geq N - \left\lceil \frac{N}{L + 1} \right\rceil \geq N - \frac{L + 2}{L + 1} \left\lfloor \frac{N}{L + 1} \right\rfloor \\ &\geq N - \frac{L + 2}{L + 1} \frac{N}{L + 1} \\ &= \frac{N}{(L + 1)^2} ((L + 1)^2 - L - 2) \geq \frac{N}{(L + 1)^2} L^2 \geq L^2, \end{aligned} \tag{4.7}$$

and so the first term in the final expression of (4.5) is subject to the concentration inequality from (4.4). That is,

$$\begin{aligned} &\mathbb{P} \left(\left| \frac{\log Z_{\bar{\mathcal{P}}_{n_L, L}}(\beta)}{N} - \sum_{\ell=1}^L \frac{\mathbb{E} \log Z_{\bar{n}_\ell - \bar{n}_{\ell-1}}(\beta_\ell)}{N} \right| > u \right) \\ &= \mathbb{P} \left(\left| \frac{\log Z_{\bar{\mathcal{P}}_{n_L, L}}(\beta)}{n_L} - \sum_{\ell=1}^L \frac{\mathbb{E} \log Z_{n_\ell - n_{\ell-1}}(\beta_\ell)}{n_L} \right| > \frac{N}{n_L} u \right) \\ &\leq 2L(2n_L + 1)^d \exp \left(- \frac{n_L \left(\frac{N}{n_L} u \right)^2}{18L^2 \beta_{\max}^2} \right) \\ &\stackrel{(4.7)}{\leq} 2L(2N + 1)^d \exp \left(- \frac{N \left(\frac{N}{n_L} u \right)^2}{18(L + 1)^2 \beta_{\max}^2} \right). \end{aligned} \tag{4.8}$$

We are now left with the task of controlling the second term in the final expression of (4.5).

Since all variables in \mathbf{g} are i.i.d., we have the following equality in law with respect to \mathbb{P} :

$$B(x) \stackrel{\text{dist}}{=} Z_{N-n_L}(\beta_{L+1}) \text{ for each } x \in D_{n_L}. \tag{4.9}$$

In particular, $\mathbb{E} \log B(x)$ is constant among such x , and so taking this constant as the value of y , we conclude the following from Lemma 4.2 with $f(t) = \log t$ and X having probability distribution given by $A(\cdot)/A$. If

$$\left| \log \sum_{x \in D_{n_L}} \frac{A(x)}{A} B(x) - \sum_{x \in D_{n_L}} \frac{A(x)}{A} \log B(x) \right| > 2u,$$

then

$$|\log B(x) - \mathbb{E} \log B(x)| > u \quad \text{for some } x \in D_{n_L}.$$

Also note that the following holds for all $\ell \in \llbracket 1, L+1 \rrbracket$, in particular $\ell = L+1$:

$$n_\ell - n_{\ell-1} \geq \left\lfloor \frac{N}{L+1} \right\rfloor \stackrel{(4.6)}{\geq} \frac{L+1}{L+2} \left\lceil \frac{N}{L+1} \right\rceil \geq \frac{N}{L+2} \geq \frac{N}{(L+1)^2}.$$

Consequently,

$$\begin{aligned} & \mathbb{P} \left(\frac{1}{N} \left| \log \sum_{x \in D_{n_L}} \frac{A(x)}{A} B(x) - \sum_{x \in D_{n_L}} \frac{A(x)}{A} \log B(x) \right| > 2u \right) \\ & \leq \mathbb{P} \left(\bigcup_{x \in D_{n_L}} \left\{ \frac{|\log B(x) - \mathbb{E} \log B(x)|}{N} > u \right\} \right) \\ & \stackrel{(4.9), (1.4)}{\leq} |D_{n_L}| \exp \left(- \frac{(N - n_L) \left(u \frac{N}{N - n_L} \right)^2}{2\beta_{L+1}^2} \right) \\ & \leq (2N + 1)^d \exp \left(- \frac{N \left(u \frac{N}{N - n_L} \right)^2}{2(L+1)^2 \beta_{L+1}^2} \right). \end{aligned} \tag{4.10}$$

Moreover, by writing

$$\left| \mathbb{E} \log Z_{N-n_L}(\beta_{L+1}) - \sum_{x \in D_{n_L}} \frac{A(x)}{A} \log B(x) \right| \leq \sum_{x \in D_{n_L}} \frac{A(x)}{A} |\mathbb{E} \log B(x) - \log B(x)|,$$

we can repeat the previous estimate to obtain

$$\begin{aligned} & \mathbb{P} \left(\frac{1}{N} \left| \mathbb{E} \log Z_{N-n_L}(\beta_{L+1}) - \sum_{x \in D_{n_L}} \frac{A(x)}{A} \log B(x) \right| > u \right) \\ & \leq \mathbb{P} \left(\bigcup_{x \in D_{n_L}} \left\{ \frac{|\log B(x) - \mathbb{E} \log B(x)|}{N} > u \right\} \right) \\ & \leq (2N + 1)^d \exp \left(- \frac{N \left(u \frac{N}{N - n_L} \right)^2}{2(L+1)^2 \beta_{L+1}^2} \right). \end{aligned} \tag{4.11}$$

Together, (4.10) and (4.11) yield

$$\begin{aligned} & \mathbb{P} \left(\frac{1}{N} \left| \mathbb{E} \log Z_{N-n_L}(\beta_{L+1}) - \log \sum_{x \in D_{n_L}} \frac{A(x)}{A} B(x) \right| > 3u \right) \\ & \leq 2(2N + 1)^d \exp \left(- \frac{N \left(u \frac{N}{N - n_L} \right)^2}{2(L+1)^2 \beta_{L+1}^2} \right). \end{aligned} \tag{4.12}$$

Putting together (4.5), (4.8), and (4.12), we conclude

$$\begin{aligned} & \mathbb{P}\left(\left|\frac{\log Z_{\mathcal{P}_{N,L+1}}(\boldsymbol{\beta}, \beta_{L+1})}{N} - \sum_{\ell=1}^{L+1} \frac{\mathbb{E} \log Z_{n_\ell - n_{\ell-1}}(\beta_\ell)}{N}\right| > u\right) \\ & \leq \mathbb{P}\left(\left|\frac{\log Z_{\mathcal{P}_{n_L,L}}(\boldsymbol{\beta})}{N} - \sum_{\ell=1}^L \frac{\mathbb{E} \log Z_{\bar{n}_\ell - \bar{n}_{\ell-1}}(\beta_\ell)}{N}\right| > u \frac{n_L}{N}\right) \\ & \quad + \mathbb{P}\left(\frac{1}{N} \left| \mathbb{E} \log Z_{N-n_L}(\beta_{L+1}) - \log \sum_{x \in D_{n_L}} \frac{A(x)}{A} B(x) \right| > u \frac{N-n_L}{N}\right) \\ & \leq 2L(2N+1)^d \exp\left(-\frac{Nu^2}{18(L+1)^2 \beta_{\max}^2}\right) + 2(2N+1)^d \exp\left(-\frac{N(\frac{u}{3})^2}{2(L+1)^2 \beta_{L+1}^2}\right) \\ & = 2(L+1)(2N+1)^d \exp\left(-\frac{Nu^2}{18(L+1)^2 (\beta_{\max} \vee \beta_{L+1})^2}\right), \end{aligned}$$

which verifies the inductive step needed for (4.4). □

5 Proof of Theorem 3.1

In preparation for the proof, we introduce the main input, Theorem D, from [7]. The statement is exactly the same as Theorem C but holds for more general Gaussian disordered systems. So that there is no confusion caused by duplicate notation, let us introduce a generic setting.

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be an abstract probability space, and $(\Sigma_N)_{N \geq 1}$ a sequence of Polish spaces equipped respectively with probability measures $(\nu_N)_{N \geq 1}$. For each N , we consider a centered Gaussian field $(\mathcal{H}_N(\sigma))_{\sigma \in \Sigma_N}$ defined on Ω . Regarding this field as a Hamiltonian, we denote the associated Gibbs measure by

$$\nu_{N,\beta}(\mathrm{d}\sigma) := \frac{e^{\beta \mathcal{H}_N(\sigma)}}{\mathcal{Z}_N(\beta)} \nu_N(\mathrm{d}\sigma), \quad \text{where } \mathcal{Z}_N(\beta) := \int_{\Sigma_N} e^{\beta \mathcal{H}_N(\sigma)} \nu_N(\mathrm{d}\sigma). \quad (5.1)$$

We make the following assumptions:

- There is a deterministic function $\mathcal{P}: [0, \infty) \rightarrow \mathbb{R}$ and a deterministic sequence $(a_N)_{N \geq 1}$ tending to infinity,² such that

$$\lim_{N \rightarrow \infty} \frac{\log \mathcal{Z}_N(\beta)}{a_N} = \mathcal{P}(\beta) \quad \mathbf{P}\text{-a.s. and in } L^1(\mathbf{P}), \text{ for every } \beta \geq 0. \quad (\text{A1})$$

- For every $\sigma \in \Sigma_N$, we have

$$\text{Var } \mathcal{H}_N(\sigma) = a_N. \quad (\text{A2})$$

- For any $\sigma^1, \sigma^2 \in \Sigma_N$, we have

$$\mathcal{R}_N(\sigma^1, \sigma^2) := \text{Corr}(\mathcal{H}_N(\sigma^1), \mathcal{H}_N(\sigma^2)) = \frac{1}{a_N} \text{Cov}(\mathcal{H}_N(\sigma^1), \mathcal{H}_N(\sigma^2)) \geq 0. \quad (\text{A3})$$

²Strictly speaking, [7] considers only the case $a_N = N$, although this is just for purposes of exposition. Even so, this single case would be enough for our purposes, since we will ultimately apply Theorem D with $a_N = n_\ell(N) - n_{\ell-1}(N)$. Indeed, the associated sequence of partitions $(\mathcal{P}_{N,L})_{N \geq 1}$ from (4.1) is contained in the union of finitely many sequences of the form $(\mathcal{P}'_{N_i,L})_{i \geq 1}$, where $n'_\ell(N_i) - n'_{\ell-1}(N_i) = i$. Therefore, one can safely apply Theorem D along each one of these sequences, and then the proof of Corollary 5.2 goes through by choosing the maximum J , maximum N_* , and minimum δ resulting from these applications.

- For each N , there exist measurable real-valued functions $(\varphi_{i,N})_{i=1}^\infty$ on Σ_N and i.i.d. standard normal random variables $(g_{i,N})_{i=1}^\infty$ defined on Ω such that for each $\sigma \in \Sigma_N$,

$$\mathcal{H}_N(\sigma) = \sum_{i=1}^\infty g_{i,N} \varphi_{i,N}(\sigma) \quad \mathbf{P}\text{-a.s.}, \tag{A4}$$

where the series on the right converges in $L^2(\mathbf{P})$.

Theorem D. [7, Thm. 1.2] Assume (A1)–(A4). If $\beta > 0$ is a point of differentiability for $\mathcal{P}(\cdot)$ with $\mathcal{P}'(\beta) < \beta$, then for every $\varepsilon > 0$, there exist integers $J = J(\beta, \varepsilon)$ and $N_* = N_*(\beta, \varepsilon)$ and a number $\delta = \delta(\beta, \varepsilon) > 0$ such that the following is true for all $N \geq N_*$. With \mathbf{P} -probability at least $1 - \varepsilon$, there exist $\sigma^1, \dots, \sigma^J \in \Sigma_N$ such that

$$\nu_{N,\beta} \left(\bigcup_{j=1}^J \{ \sigma \in \Sigma_N : R(\sigma^j, \sigma) \geq \delta \} \right) \geq 1 - \varepsilon.$$

Returning to the polymer setting, we consider the following modifications to the random environment: restricting to times inside the interval $\llbracket n_{\ell-1} + 1, n_\ell \rrbracket$, and restricting to times outside the interval. The resulting Hamiltonians will be written as

$$\begin{aligned} H_N^{(\ell)}(\sigma) &:= \sum_{i \in \llbracket n_{\ell-1} + 1, n_\ell \rrbracket} g_N(i, \sigma_i), \\ \widehat{H}_N^{(\ell)}(\sigma) &:= \sum_{i \in \llbracket 1, N \rrbracket \setminus \llbracket n_{\ell-1} + 1, n_\ell \rrbracket} g_N(i, \sigma_i), \end{aligned}$$

where $(g_N(i, x) : N \geq 1, i \in \llbracket 1, N \rrbracket, x \in \mathbb{Z}^d)$ is an i.i.d. collection even across N (recall Remark 1.1). Therefore, we can partition g into the following pair of independent sub-collections,

$$\begin{aligned} g^{(\ell)} &:= \{g_N(i, x) : N \geq 1, i \in \llbracket n_{\ell-1} + 1, n_\ell \rrbracket, x \in \mathbb{Z}^d\}, \\ \widehat{g}^{(\ell)} &:= \{g_N(i, x) : N \geq 1, i \in \llbracket 1, N \rrbracket \setminus \llbracket n_{\ell-1} + 1, n_\ell \rrbracket, x \in \mathbb{Z}^d\}, \end{aligned}$$

and then $H_N^{(\ell)}(\cdot)$ is a function of $g^{(\ell)}$, while $\widehat{H}_N^{(\ell)}(\cdot)$ is a function of $\widehat{g}^{(\ell)}$. We will write $\mathbb{P}_{\widehat{g}^{(\ell)}}$ to denote the probability measure obtained by conditioning \mathbb{P} on $\widehat{g}^{(\ell)}$, and $\mathbb{E}_{\widehat{g}^{(\ell)}}$ will denote expectation with respect to $\mathbb{P}_{\widehat{g}^{(\ell)}}$ (i.e. integrating over just $g^{(\ell)}$). While the law of $g^{(\ell)}$ is no different under $\mathbb{P}_{\widehat{g}^{(\ell)}}$ than under \mathbb{P} , these notational devices will make clearer how we invoke Theorem D and avoid the slightly more cumbersome $\mathbb{P}(\cdot | \widehat{g}^{(\ell)})$.

Next, for each $\ell \in \llbracket 1, L \rrbracket$ we introduce the following probability measure on Σ :

$$\widehat{\mu}_{N,\beta}^{(\ell)}(d\sigma) = \frac{1}{\widehat{Z}_N^{(\ell)}(\beta)} e^{\beta \widehat{H}_N^{(\ell)}(\sigma)} P(d\sigma), \quad \widehat{Z}_N^{(\ell)}(\beta) := E(e^{\beta \widehat{H}_N^{(\ell)}(\sigma)}).$$

Observe that

$$\begin{aligned} \mu_{N,\beta}(d\sigma) &= \frac{1}{Z_N(\beta)} e^{\beta H_N^{(\ell)}(\sigma)} e^{\beta \widehat{H}_N^{(\ell)}(\sigma)} P(d\sigma) \\ &= \frac{\widehat{Z}_N^{(\ell)}(\beta)}{Z_N(\beta)} e^{\beta H_N^{(\ell)}(\sigma)} \widehat{\mu}_{N,\beta}^{(\ell)}(d\sigma) = \frac{1}{Z_N^{(\ell)}(\beta)} e^{\beta H_N^{(\ell)}(\sigma)} \widehat{\mu}_{N,\beta}^{(\ell)}(d\sigma), \end{aligned} \tag{5.2}$$

where

$$Z_N^{(\ell)}(\beta) := \frac{Z_N(\beta)}{\widehat{Z}_N^{(\ell)}(\beta)}. \tag{5.3}$$

We will ultimately apply Theorem D to this “restricted” setting, using

$$\begin{aligned} \Sigma_N = \Sigma \text{ from (1.1),} \quad \nu_N = \widehat{\mu}_{N,\beta}^{(\ell)}, \quad \mathbf{P} = \mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}, \quad \mathcal{H}_N = H_N^{(\ell)}, \quad (5.4) \\ a_N = n_\ell - n_{\ell-1}, \quad \nu_{N,\beta} = \mu_{N,\beta}, \quad \mathcal{Z}_N(\beta) = Z_N^{(\ell)}(\beta), \quad \mathcal{R}_N(\cdot, \cdot) = R^{(\ell)}(\cdot, \cdot). \end{aligned}$$

Note that ν_N and \mathbf{P} are now random measures depending on $\widehat{\mathbf{g}}^{(\ell)}$, but since the Hamiltonian \mathcal{H}_N is independent of this randomness, Theorem D will still apply. With these choices, we first need to verify (A1).

Proposition 5.1. *For each $\ell \in \llbracket 1, L \rrbracket$ and any $\beta \geq 0$, the following statement holds \mathbb{P} -almost surely:*

$$\lim_{N \rightarrow \infty} \frac{\log Z_N^{(\ell)}(\beta)}{n_\ell - n_{\ell-1}} = p(\beta) \quad \mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}\text{-a.s. and in } L^\alpha(\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}) \text{ for all } \alpha \in [1, \infty). \quad (5.5)$$

Proof. If $\beta = 0$, then $Z_N^{(\ell)}(0) = 1$ is deterministic, and so (5.5) holds trivially with $p(0) = 0$. Consequently, we may assume $\beta > 0$.

By Theorem 4.1, we know

$$\lim_{N \rightarrow \infty} \frac{\log Z_N(\beta)}{N} = p(\beta), \quad \lim_{N \rightarrow \infty} \frac{\log \widehat{Z}_N^{(\ell)}(\beta)}{N} = \frac{L-1}{L} p(\beta), \quad (5.6)$$

where the limits are \mathbb{P} -almost sure and in $L^\alpha(\mathbb{P})$ for every $\alpha \in [1, \infty)$. By definition (5.3) and the fact that $(n_\ell - n_{\ell-1}) \sim N/L$, the following limit thus holds in the same senses:

$$\lim_{N \rightarrow \infty} \frac{\log Z_N^{(\ell)}(\beta)}{n_\ell - n_{\ell-1}} = p(\beta). \quad (5.7)$$

In particular, since (5.7) holds \mathbb{P} -almost surely, Fubini’s theorem guarantees the following: For almost every realization of $\widehat{\mathbf{g}}^{(\ell)}$, (5.7) holds $\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}$ -almost surely. This proves the first part of (5.5).

Meanwhile, for any $\varepsilon > 0$ and $\alpha \geq 1$, Markov’s inequality gives

$$\begin{aligned} \mathbb{P}\left(\mathbb{E}_{\widehat{\mathbf{g}}^{(\ell)}} \left| \frac{\log Z_N(\beta)}{N} - \frac{\mathbb{E} \log Z_N(\beta)}{N} \right|^\alpha > \varepsilon\right) &\leq \varepsilon^{-1} \mathbb{E} \left| \frac{\log Z_N(\beta)}{N} - \frac{\mathbb{E} \log Z_N(\beta)}{N} \right|^\alpha \\ &= \varepsilon^{-1} \int_0^\infty \mathbb{P}\left(\left| \frac{\log Z_N(\beta)}{N} - \frac{\mathbb{E} \log Z_N(\beta)}{N} \right|^\alpha > u\right) du \\ &\stackrel{(1.4)}{\leq} \varepsilon^{-1} \int_0^\infty \exp\left(-\frac{Nu^{2/\alpha}}{2\beta^2}\right) du = CN^{-\alpha/2}, \end{aligned}$$

where C depends on α and β but not on N . By taking $\alpha > 2$, we can apply Borel–Cantelli to determine that with \mathbb{P} -probability one,

$$\limsup_{N \rightarrow \infty} \mathbb{E}_{\widehat{\mathbf{g}}^{(\ell)}} \left| \frac{\log Z_N(\beta)}{N} - \frac{\mathbb{E} \log Z_N(\beta)}{N} \right|^\alpha \leq \varepsilon.$$

By taking a countable sequence $\varepsilon_k \searrow 0$, we further deduce

$$\lim_{N \rightarrow \infty} \mathbb{E}_{\widehat{\mathbf{g}}^{(\ell)}} \left| \frac{\log Z_N(\beta)}{N} - \frac{\mathbb{E} \log Z_N(\beta)}{N} \right|^\alpha = 0 \quad \mathbb{P}\text{-a.s.} \quad (5.8)$$

Since (1.3) gives the deterministic limit $\frac{1}{N} \mathbb{E} \log Z_N(\beta) \rightarrow p(\beta)$, it follows from (5.8) that with \mathbb{P} -probability one we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log Z_N(\beta) = p(\beta) \quad \text{in } L^\alpha(\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}). \quad (5.9a)$$

Moreover, given that α can be taken arbitrarily large, this convergence occurs in $L^\alpha(\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}})$ simultaneously for all $\alpha \in [1, \infty)$. On the other hand, from (5.6) we know

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \widehat{Z}_N^{(\ell)}(\beta) = \frac{L-1}{L} p(\beta), \tag{5.9b}$$

also with \mathbb{P} -probability one. Furthermore, since $\widehat{Z}_n^{(\ell)}(\beta)$ is determined entirely by $\widehat{\mathbf{g}}^{(\ell)}$, this last limit is a deterministic statement with respect to $\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}$; in particular, it holds in $L^\alpha(\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}})$. From (5.3), (5.9), and the fact that $(n_\ell - n_{\ell-1}) \sim N/L$, we now have

$$\lim_{N \rightarrow \infty} \mathbb{E}_{\widehat{\mathbf{g}}^{(\ell)}} \left| \frac{\log Z_N^{(\ell)}(\beta)}{n_\ell - n_{\ell-1}} - p(\beta) \right|^\alpha = 0 \quad \mathbb{P}\text{-a.s.}$$

We have thus verified both parts of (5.5). □

Given Proposition 5.1, we can make a statement approaching Theorem 3.1. The following result asserts that once the system size becomes large enough, the “external” disorder $\widehat{\mathbf{g}}^{(\ell)}$ becomes sufficiently well behaved so that when only the “internal” disorder $\mathbf{g}^{(\ell)}$ is regarded as random, the polymer along the subinterval $[[n_{\ell-1} + 1, n_\ell]]$ admits the same localization statement as in Theorem C.

Corollary 5.2. *Let $\ell \in \llbracket 1, L \rrbracket$, and assume $\beta > 0$ is a point of differentiability for $p(\cdot)$ with $p'(\beta) < \beta$. Then with \mathbb{P} -probability one, the following is true. For every $\varepsilon > 0$, there exist integers $J = J(\beta, \varepsilon, \widehat{\mathbf{g}}^{(\ell)})$ and $N_* = N_*(\beta, \varepsilon, \widehat{\mathbf{g}}^{(\ell)})$ and a number $\delta = \delta(\beta, \varepsilon, \widehat{\mathbf{g}}^{(\ell)}) > 0$ such that for all $N \geq N_*$, the following event has $\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}$ -probability at least $1 - \varepsilon$:*

$$\mathbb{E}_{N,J,\delta,\varepsilon}^{(\ell)} := \left\{ \exists \sigma^1, \dots, \sigma^J \in \Sigma : \mu_{N,\beta} \left(\bigcup_{j=1}^J \{ \sigma \in \Sigma : R^{(\ell)}(\sigma^j, \sigma) \geq \delta \} \right) \geq 1 - \varepsilon \right\}.$$

Proof. Because $\widehat{\mathbf{g}}^{(\ell)}$ and $\mathbf{g}^{(\ell)}$ are independent, the law of the latter given the former remains i.i.d. standard normal. Therefore, (5.2) is a representation of $\mu_{N,\beta}$ in the form of (5.1). Proposition 5.1 verifies that in this representation, the assumption (A1) holds. Also, it is trivial to check that

$$\text{Cov}(H_N^{(\ell)}(\sigma^1), H_N^{(\ell)}(\sigma^2)) = (n_\ell - n_{\ell-1}) R^{(\ell)}(\sigma^1, \sigma^2) \geq 0, \quad \sigma^1, \sigma^2 \in \Sigma.$$

In particular, we have

$$\text{Var } H_N^{(\ell)}(\sigma) = n_\ell - n_{\ell-1}, \quad \sigma \in \Sigma. \tag{5.10}$$

Thus (A2)–(A4) also hold, and we can apply Theorem D with the identifications in (5.4) to obtain the result. □

Recall that \mathbf{g} is supported on a probability space we denote $(\Omega, \mathcal{F}, \mathbb{P})$. Following Corollary 5.2, we can define the event

$$\widehat{\mathbb{E}}_{N,J,\delta,\varepsilon}^{(\ell)} := \{ \mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}(\mathbb{E}_{N,J,\delta,\varepsilon}^{(\ell)}) \geq 1 - \varepsilon \}. \tag{5.11}$$

Let us postpone verification that such an event is measurable, and proceed directly to the proof of Theorem 3.1.

Proof of Theorem 3.1. Let $\varepsilon > 0$ and L be given. In the notation of Corollary 5.2, it suffices to find J, N_* , and δ such that

$$\mathbb{P} \left(\bigcap_{\ell=1}^L \mathbb{E}_{N,J,\delta,\varepsilon/L}^{(\ell)} \right) \geq 1 - \varepsilon \quad \text{for all } N \geq N_*, \tag{5.12}$$

since the event $\bigcap_{\ell=1}^L E_{N,J,\delta,\varepsilon/L}^{(\ell)}$ implies the existence of $\sigma^1, \dots, \sigma^{JL} \in \Sigma$ such that

$$\mu_{N,\beta} \left(\bigcap_{\ell=1}^L \bigcup_{j=1}^{JL} \{ \sigma \in \Sigma : R^{(\ell)}(\sigma^j, \sigma) \geq \delta \} \right) \geq 1 - \varepsilon.$$

The conclusion of Theorem 3.1 then follows by replacing J with JL . The remainder of the proof is thus establishing (5.12).

Let $\varepsilon' = \varepsilon'(\varepsilon, L)$ be a positive number to be specified later. Take $(\delta_k)_{k \geq 1}$ to be any decreasing sequence tending to 0 as $k \rightarrow \infty$. From Corollary 5.2, we know

$$\mathbb{P} \left(\bigcup_{J=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcup_{N_*=1}^{\infty} \bigcap_{N=N_*}^{\infty} \widehat{E}_{N,J,\delta_k,\varepsilon'}^{(\ell)} \right) = 1 \quad \text{for each } \ell \in \llbracket 1, L \rrbracket.$$

Since $\widehat{E}_{N,J,\delta_k,\varepsilon'} \subset \widehat{E}_{N,J+1,\delta_k,\varepsilon'}$, we can choose $J = J(\beta, \varepsilon', L)$ sufficiently large that

$$\mathbb{P} \left(\bigcup_{k=1}^{\infty} \bigcup_{N_*=1}^{\infty} \bigcap_{N=N_*}^{\infty} \widehat{E}_{N,J,\delta_k,\varepsilon'}^{(\ell)} \right) \geq 1 - \varepsilon' \quad \text{for each } \ell \in \llbracket 1, L \rrbracket.$$

By the assumption $\delta_k > \delta_{k+1}$, we also have $\widehat{E}_{N,J,\delta_k,\varepsilon'} \subset \widehat{E}_{N,J,\delta_{k+1},\varepsilon'}$, and so we can choose $k = k(\beta, \varepsilon', L, J)$ sufficiently large that

$$\mathbb{P} \left(\bigcup_{N_*=1}^{\infty} \bigcap_{N=N_*}^{\infty} \widehat{E}_{N,J,\delta_k,\varepsilon'}^{(\ell)} \right) \geq 1 - 2\varepsilon' \quad \text{for each } \ell \in \llbracket 1, L \rrbracket.$$

Henceforth we simply write $\delta = \delta_k$. Finally, we choose $N_* = N_*(\beta, \varepsilon', L, J, \delta)$ sufficiently large that

$$\mathbb{P} \left(\bigcap_{N=N_*}^{\infty} \widehat{E}_{N,J,\delta,\varepsilon'}^{(\ell)} \right) \geq 1 - 3\varepsilon' \quad \text{for each } \ell \in \llbracket 1, L \rrbracket.$$

Now, whether or not $\widehat{E}_{N,J,\delta,\varepsilon'}^{(\ell)}$ occurs depends only on $\widehat{\mathbf{g}}^{(\ell)}$; see Proposition 5.3. By definition (5.11), when $\widehat{E}_{N,J,\delta,\varepsilon'}^{(\ell)}$ does occur, the event $E_{N,J,\delta,\varepsilon'}^{(\ell)}$ has $\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}$ -probability at least $1 - \varepsilon'$. Therefore, by our choice of N_* , we have

$$\begin{aligned} \mathbb{P}(E_{N,J,\delta,\varepsilon'}^{(\ell)}) &= \mathbb{E}(\mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}(E_{N,J,\delta,\varepsilon'}^{(\ell)})) \\ &\geq \mathbb{E}(\mathbf{1}_{\widehat{E}_{N,J,\delta,\varepsilon'}^{(\ell)}} \mathbb{P}_{\widehat{\mathbf{g}}^{(\ell)}}(E_{N,J,\delta,\varepsilon'}^{(\ell)})) \\ &\geq (1 - 3\varepsilon')(1 - \varepsilon') \geq 1 - 4\varepsilon' \quad \text{for each } \ell \in \llbracket 1, L \rrbracket, N \geq N_*, \end{aligned}$$

and thus

$$\mathbb{P} \left(\bigcap_{\ell=1}^L E_{N,J,\delta,\varepsilon'}^{(\ell)} \right) \geq 1 - 4L\varepsilon' \quad \text{for all } N \geq N_*.$$

To complete the proof, we set $\varepsilon' = \varepsilon/(4L)$ and note that

$$\mathbb{P} \left(\bigcap_{\ell=1}^L E_{N,J,\delta,\varepsilon}^{(\ell)} \right) \geq \mathbb{P} \left(\bigcap_{\ell=1}^L E_{N,J,\delta,\varepsilon'}^{(\ell)} \right) \geq 1 - \varepsilon \quad \text{for all } N \geq N_*. \quad \square$$

To conclude the section, we return to the technical issue of measurability for the event $\widehat{E}_{N,J,\delta,\varepsilon}^{(\ell)}$ defined in (5.11). Let $\widehat{\mathcal{F}}^{(\ell)}$ denote the sub-sigma-algebra of \mathcal{F} generated by $\widehat{\mathbf{g}}^{(\ell)}$.

Proposition 5.3. For any integers $N \geq L$, $J \geq 1$, $\ell \in \llbracket 1, L \rrbracket$, and numbers $\delta, \varepsilon > 0$, we have $\widehat{E}_{N,J,\delta,\varepsilon}^{(\ell)} \in \widehat{\mathcal{F}}^{(\ell)}$.

We will make use of two lemmas.

Lemma 5.4. If r is a non-constant rational function in n real variables, then for any $t \in \mathbb{R}$, the set $\{x \in \mathbb{R}^n : r(x) = t\}$ has zero Lebesgue measure.

Proof. Let us write $r = f/g$, where f and g are polynomials. Then $r - t = (f - tg)/g$, which vanishes if and only if $f - tg = 0$. By hypothesis, $f - tg$ is not identically equal to 0, and so this polynomial may only vanish on a set of Lebesgue measure zero [21]. \square

Lemma 5.5. Let $\mathbf{X} \in \mathbb{R}^n$ be a random vector supported on $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $f: \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ is a continuous function such that

$$\mathbb{P}(f(\mathbf{X}, \mathbf{y}) = t) = 0 \quad \text{for any } \mathbf{y} \in \mathbb{R}^m, t \in \mathbb{R}.$$

Then the map $\mathbf{y} \mapsto \mathbb{P}(f(\mathbf{X}, \mathbf{y}) \geq t)$ is continuous from \mathbb{R}^m to $[0, 1]$, for any $t \in \mathbb{R}$.

Proof. Fix $\mathbf{y} \in \mathbb{R}^m$ and $t \in \mathbb{R}$, and let $\varepsilon > 0$ be given. By hypothesis, we can choose $h > 0$ so small that $\mathbb{P}(f(\mathbf{X}, \mathbf{y}) \in [t - h, t + h]) < \varepsilon$. Next choose $K \subset \mathbb{R}^n$ to be a compact set sufficiently large that $\mathbb{P}(\mathbf{X} \notin K) < \varepsilon$. By uniform continuity of f on $K \times (\mathbf{y} + [-1, 1]^m)$, we may choose $\delta > 0$ sufficiently small that

$$\mathbf{X} \in K, \delta \in \mathbb{R}^m, \|\delta\| < \delta \implies |f(\mathbf{X}, \mathbf{y} + \delta) - f(\mathbf{X}, \mathbf{y})| \leq h.$$

It now follows that whenever $\|\delta\| < \delta$, we have

$$\mathbb{P}(f(\mathbf{X}, \mathbf{y} + \delta) \geq t) \geq \mathbb{P}(\mathbf{X} \in K, f(\mathbf{X}, \mathbf{y}) \geq t + h) \geq \mathbb{P}(f(\mathbf{X}, \mathbf{y}) \geq t) - 2\varepsilon,$$

as well as

$$\mathbb{P}(f(\mathbf{X}, \mathbf{y} + \delta) < t) \geq \mathbb{P}(\mathbf{X} \in K, f(\mathbf{X}, \mathbf{y}) < t - h) \geq \mathbb{P}(f(\mathbf{X}, \mathbf{y}) < t) - 2\varepsilon.$$

The two previous displays together imply

$$|\mathbb{P}(f(\mathbf{X}, \mathbf{y} + \delta) \geq t) - \mathbb{P}(f(\mathbf{X}, \mathbf{y}) \geq t)| \leq 2\varepsilon.$$

As ε is arbitrary, the continuity of the map $\mathbf{y} \mapsto \mathbb{P}(f(\mathbf{X}, \mathbf{y}) \geq t)$ has been proved. \square

Proof of Proposition 5.3. Recall the set $D_i = \{x \in \mathbb{Z}^d : P(\sigma_i = x) > 0\}$, for $i \geq 1$. For every finite N , the random Hamiltonian $H_N(\cdot)$ depends only on a finite set of random variables, namely

$$\mathbf{g}_N := \{g_N(i, x) : 1 \leq i \leq N, x \in D_i\}.$$

As before, let us partition this collection is the following disjoint sub-collections:

$$\begin{aligned} \mathbf{g}_N^{(\ell)} &:= \{g_N(i, x) : i \in \llbracket n_{\ell-1} + 1, n_\ell \rrbracket, x \in D_i\}, \\ \widehat{\mathbf{g}}_N^{(\ell)} &:= \{g_N(i, x) : i \in \llbracket 1, N \rrbracket \setminus \llbracket n_{\ell-1} + 1, n_\ell \rrbracket, x \in D_i\}. \end{aligned}$$

We will restrict our attention to these finite collections and thus regard all subsequent statements as concerning only finite-dimensional vectors. In keeping with this finite-dimensional perspective, we will write Σ_N to denote the set of the $(2d)^N$ possible simple random walk paths starting at the origin and consisting of N steps. It is natural to regard $\mu_{N,\beta}$ and $\widehat{\mu}_{N,\beta}^{(\ell)}$ as measures on Σ_N , as opposed to Σ defined in (1.1). Similarly, it makes sense to take (2.2) as a definition of $R^{(\ell)}(\sigma^1, \sigma^2)$ for $\sigma^1, \sigma^2 \in \Sigma_N$.

Now consider subsets of Σ_N of the form

$$S_{\sigma^1, \dots, \sigma^J, \delta}^{(\ell)} := \bigcup_{j=1}^J \{ \sigma \in \Sigma_N : R^{(\ell)}(\sigma^j, \sigma) \geq \delta \}, \quad \sigma^1, \dots, \sigma^J \in \Sigma_N, \delta > 0, \ell \in \llbracket 1, L \rrbracket.$$

The event $E_{N, J, \delta, \varepsilon}^{(\ell)}$ from Corollary 5.2 can be expressed as

$$\begin{aligned} E_{N, J, \delta, \varepsilon}^{(\ell)} &= \bigcup_{\sigma^1, \dots, \sigma^J \in \Sigma_N} \{ \mu_{N, \beta}(S_{\sigma^1, \dots, \sigma^J, \delta}) \geq 1 - \varepsilon \} \\ &= \left\{ \max_{\sigma^1, \dots, \sigma^J \in \Sigma_N} \mu_{N, \beta}(S_{\sigma^1, \dots, \sigma^J, \delta}) \geq 1 - \varepsilon \right\} \\ &= \bigcap_{n=1}^{\infty} \bigcup_{q=1}^{\infty} \left\{ \left(\frac{1}{|\Sigma_N|^J} \sum_{\sigma^1, \dots, \sigma^J \in \Sigma_N} \mu_{N, \beta}(S_{\sigma^1, \dots, \sigma^J, \delta})^q \right)^{1/q} \geq 1 - \varepsilon - \frac{1}{n} \right\}. \end{aligned}$$

By monotonicity, it is evident that

$$\begin{aligned} &\mathbb{P}_{\hat{g}^{(\ell)}}(E_{N, J, \delta, \varepsilon}^{(\ell)}) \\ &= \lim_{n \rightarrow \infty} \lim_{q \rightarrow \infty} \mathbb{P}_{\hat{g}^{(\ell)}} \left(\left(\frac{1}{|\Sigma_N|^J} \sum_{\sigma^1, \dots, \sigma^J \in \Sigma_N} \mu_{N, \beta}(S_{\sigma^1, \dots, \sigma^J, \delta})^q \right)^{1/q} \geq 1 - \varepsilon - \frac{1}{n} \right). \end{aligned} \quad (5.13)$$

Now we observe that the quantity

$$\frac{1}{|\Sigma_N|^J} \sum_{\sigma^1, \dots, \sigma^J \in \Sigma_N} \mu_{N, \beta}(S_{\sigma^1, \dots, \sigma^J, \delta})^q$$

is a non-constant rational function in the variables $\{e^{g(i, x)} : i \geq 1, x \in D_i\}$. Moreover, it remains non-constant even if the realization of $\hat{g}_N^{(\ell)}$ is fixed (assuming $n_{\ell-1}(N) < n_{\ell}(N)$, which is true so long as $N \geq L$). Therefore, Lemma 5.4 tells us that the map

$$(\mathbf{g}_N^{(\ell)}, \hat{\mathbf{g}}_N^{(\ell)}) \mapsto \left(\frac{1}{|\Sigma_N|^J} \sum_{\sigma^1, \dots, \sigma^J \in \Sigma_N} \mu_{N, \beta}(S_{\sigma^1, \dots, \sigma^J, \delta})^q \right)^{1/q}$$

satisfies the hypotheses of Lemma 5.5 (note that $e^{g(i, x)}$ is a continuous function of $g(i, x)$ and has a density with respect to Lebesgue measure). In turn, Lemma 5.5 guarantees the continuity—in particular, measurability—of the map

$$\hat{\mathbf{g}}^{(\ell)} \mapsto \mathbb{P}_{\hat{\mathbf{g}}^{(\ell)}} \left(\left(\frac{1}{|\Sigma_N|^J} \sum_{\sigma^1, \dots, \sigma^J \in \Sigma_N} \mu_{N, \beta}(S_{\sigma^1, \dots, \sigma^J, \delta})^q \right)^{1/q} \geq 1 - \varepsilon - \frac{1}{n} \right).$$

Consequently, (5.13) exhibits $\hat{\mathbf{g}}^{(\ell)} \mapsto \mathbb{P}_{\hat{\mathbf{g}}^{(\ell)}}(E_{N, J, \delta, \varepsilon}^{(\ell)})$ as a limit of measurable functions, implying that this map is also measurable. In particular, the event $\hat{E}_{N, J, \delta, \varepsilon}^{(\ell)}$ considered in (5.11) is $\hat{\mathcal{F}}^{(\ell)}$ -measurable. \square

References

- [1] Aizenman, M., Lebowitz, J. L., and Ruelle, D. Some rigorous results on the Sherrington-Kirkpatrick spin glass model. *Comm. Math. Phys.* 112, 1 (1987), 3–20. MR-0904135
- [2] Anderson, P. W. Absence of diffusion in certain random lattices. *Phys. Rev.* 109 (Mar 1958), 1492–1505.
- [3] Auffinger, A., and Louidor, O. Directed polymers in a random environment with heavy tails. *Comm. Pure Appl. Math.* 64, 2 (2011), 183–204. MR-2766526

- [4] Bakhtin, Y., and Seo, D. Localization of directed polymers in continuous space. *Electron. J. Probab.* 25 (2020), Paper No. 142, 56. MR-4186261
- [5] Bates, E. Localization of directed polymers with general reference walk. *Electron. J. Probab.* 23 (2018), Paper No. 30, 45. MR-3785400
- [6] Bates, E., and Chatterjee, S. The endpoint distribution of directed polymers. *Ann. Probab.* 48, 2 (2020), 817–871. MR-4089496
- [7] Bates, E., and Chatterjee, S. Localization in Gaussian disordered systems at low temperature. *Ann. Probab.* 48, 6 (2020), 2755–2806. MR-4164453
- [8] Berger, Q., and Lacoïn, H. The Scaling Limit of the Directed Polymer with Power-Law Tail Disorder. *Comm. Math. Phys.* (2021).
- [9] Berger, Q., and Torri, N. Directed polymers in heavy-tail random environment. *Ann. Probab.* 47, 6 (2019), 4024–4076. MR-4038048
- [10] Carmona, P., and Hu, Y. On the partition function of a directed polymer in a Gaussian random environment. *Probab. Theory Related Fields* 124, 3 (2002), 431–457. MR-1939654
- [11] Chatterjee, S. Proof of the Path Localization Conjecture for Directed Polymers. *Comm. Math. Phys.* 370, 2 (2019), 703–717. MR-3994583
- [12] Comets, F. *Directed polymers in random environments*, vol. 2175 of *Lecture Notes in Mathematics*. Springer, Cham, 2017. Lecture notes from the 46th Probability Summer School held in Saint-Flour, 2016. MR-3444835
- [13] Comets, F., and Cosco, C. Brownian polymers in Poissonian environment: a survey. Preprint, available at arXiv:arXiv:1805.10899.
- [14] Comets, F., and Cranston, M. Overlaps and pathwise localization in the Anderson polymer model. *Stochastic Process. Appl.* 123, 6 (2013), 2446–2471. MR-3038513
- [15] Comets, F., and Neveu, J. The Sherrington-Kirkpatrick model of spin glasses and stochastic calculus: the high temperature case. *Comm. Math. Phys.* 166, 3 (1995), 549–564. MR-1312435
- [16] Comets, F., and Nguyen, V.-L. Localization in log-gamma polymers with boundaries. *Probab. Theory Related Fields* 166, 1-2 (2016), 429–461. MR-3547743
- [17] Comets, F., and Yoshida, N. Directed polymers in random environment are diffusive at weak disorder. *Ann. Probab.* 34, 5 (2006), 1746–1770. MR-2271480
- [18] Comets, F., and Yoshida, N. Localization transition for polymers in Poissonian medium. *Comm. Math. Phys.* 323, 1 (2013), 417–447. MR-3085670
- [19] den Hollander, F. *Random polymers*, vol. 1974 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Lectures from the 37th Probability Summer School held in Saint-Flour, 2007. MR-2504175
- [20] Dey, P. S., and Zygouras, N. High temperature limits for $(1 + 1)$ -dimensional directed polymer with heavy-tailed disorder. *Ann. Probab.* 44, 6 (2016), 4006–4048. MR-3572330
- [21] (https://math.stackexchange.com/users/822/nate_eldredge), N. E. The lebesgue measure of zero set of a polynomial function is zero. Mathematics Stack Exchange. URL:<https://math.stackexchange.com/q/1920527> (version: 2016-09-09).
- [22] Huse, D. A., and Henley, C. L. Pinning and roughening of domain walls in Ising systems due to random impurities. *Phys. Rev. Lett.* 54, 25 (1985), 2708–2711.
- [23] Imbrie, J. Z., and Spencer, T. Diffusion of directed polymers in a random environment. *J. Statist. Phys.* 52, 3-4 (1988), 609–626. MR-0968950
- [24] Panchenko, D. On differentiability of the Parisi formula. *Electron. Commun. Probab.* 13 (2008), 241–247. MR-2399285
- [25] Seppäläinen, T. Scaling for a one-dimensional directed polymer with boundary conditions. *Ann. Probab.* 40, 1 (2012), 19–73. MR-2917766
- [26] Talagrand, M. Parisi measures. *J. Funct. Anal.* 231, 2 (2006), 269–286. MR-2195333
- [27] Torri, N. Pinning model with heavy tailed disorder. *Stochastic Process. Appl.* 126, 2 (2016), 542–571. MR-3434993
- [28] Vargas, V. Strong localization and macroscopic atoms for directed polymers. *Probab. Theory Related Fields* 138, 3-4 (2007), 391–410. MR-2299713

[29] Viveros, R. Directed polymer for very heavy tailed random walks. Preprint, available at arXiv:arXiv:2003.14280.

Acknowledgments. I am indebted to Sourav Chatterjee for suggesting the consideration of a time-dependent inverse temperature. This idea was the inspiration leading to the present article. I also thank Francis Comets for valuable discussion during the workshop on “Self-interacting random walks, Polymers and Folding” held at Centre International de Rencontres Mathématiques, and Hubert Lacoin for directing my attention to [10, Sec. 7]. Finally, I am very grateful to the referees for their corrections and suggestions, which improved the exposition.