



HAL
open science

Variational representation and estimates for the free energy of a quenched charged polymer model

Julien Poisat

► **To cite this version:**

Julien Poisat. Variational representation and estimates for the free energy of a quenched charged polymer model. 2025. <hal-04965911>

HAL Id: hal-04965911

<https://hal.science/hal-04965911v1>

Preprint submitted on 25 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

VARIATIONAL REPRESENTATION AND ESTIMATES FOR THE FREE ENERGY OF A QUENCHED CHARGED POLYMER MODEL

JULIEN POISAT

ABSTRACT. Random walks with a disordered self-interaction potential may be used to model charged polymers. In this paper we consider a one-dimensional and directed version of the charged polymer model that was introduced by Derrida, Griffiths and Higgs. We prove new results for the associated quenched free energy, including a variational formula based on a quenched large deviation principle established by Birkner, Greven and den Hollander. We also take the occasion to (i) provide detailed proofs for state-of-the-art results pointing towards the existence of a freezing transition and (ii) proceed with minor corrections for two results previously obtained by the present author with Caravenna, den Hollander and P  tr  lis for the undirected model.

CONTENTS

1. Introduction	1
2. The Derrida-Griffiths-Higgs directed model	2
2.1. Definition of the model	2
2.2. Predictions and state of the art	4
2.3. Existence and variational representation of the free energy	9
2.4. High-temperature estimate on the free energy	14
2.5. Low-temperature estimate on the free energy	18
3. Corrections to the undirected model	19
4. Conclusion and perspectives	21
Appendix A. Technical estimates	22
Appendix B. High-temperature expansion of the finite-volume free energy	23
References	24

1. INTRODUCTION

This paper deals with a particular class of random polymer models called *charged polymers*. Polymers are large molecules, either natural or artificial, which result from the bonding of many constitutive units (atoms or group of atoms) called *monomers*. We shall hereafter restrict to flexible *chain-like* structures. The spatial configuration (a.k.a. *conformation*) of such macromolecules may drastically change upon a small change in parameter like the temperature (e.g. the denaturation of DNA or the collapse transition of a polymer in a poor solvent). *Random polymer models* usually refer to a class of *discrete probabilistic*

Date: February 26, 2025.

2020 Mathematics Subject Classification. 60K37, 82B41, 82B44.

Key words and phrases. charged polymer, interacting random walk, disordered system, free energy, freezing transition, collapse transition, conditional large deviation principle, annealing.

JP acknowledges the support of the ANR grant LOCAL (ANR-22-CE40-0012).

models in which (i) the possible conformations are given by a subset of *finite lattice paths* and (ii) the *polymer measure* (for a given number of monomers) is a reference measure on such paths (e.g. the simple random walk measure) tilted by a Gibbs factor (the exponential of minus the energy or Hamiltonian function), following the formalism of equilibrium statistical mechanics. We refer to [11, 13, 17, 18] for a review of some of these models.

In a *charged polymer* (or polyelectrolyte), each monomer carries an electric charge that interacts with the charges on the other monomers: monomers with alike charges repel each other while monomers with opposite charges attract each other. This leads to complex interactions, especially when the charge sequence (as it is read along the polymer chain) is itself *disordered*. Following Kantor and Kardar [19], the (free) polymer chain is modeled by the simple random walk on \mathbb{Z}^d , the charges are independent and identically distributed (i.i.d.) real-valued random variables, and interactions happen between pairs of monomers at *self-intersections* only (two-body short-range interactions). For the *annealed* (i.e. averaged out charges) version of the model, a phase transition between a *collapsed* phase (attractive interactions prevail) and an *extended* phase (repulsive interactions prevail) in the inverse temperature v.s. charge bias phase diagram has been established under exponential moment conditions on the charge distribution [2, 8]. The *quenched* (frozen charges) version however remains largely open, apart from partial results showing extended behavior of the chain under a large charge bias condition [8].

In the present paper we consider the one-dimensional *directed** version of this model, which was originally introduced by Derrida, Griffiths and Higgs [15, 16]. In that model, the random walk may only move to the right or stay put (instead of moving to the left) with equal probability, at each unit step. The aforementioned authors argued in favor of a *freezing transition* between a collapsed and an extended phase, with the help of rigorous and heuristic arguments. The main purpose of this paper is to review previous results and provide new estimates for the quenched free energy associated to this model. In particular, we show that the latter fits into a larger class of one-dimensional disordered models (including disordered pinning and copolymer models) to which the Birkner-Greenden Hollander *quenched* (or conditional) *large deviation principle* [3, 4, 5] may be applied to obtain a variational formula for the free energy. This is the content of Section 2. We also take the occasion to correct two propositions from [8, Appendix D] for the *undirected* model, see Section 3.

Notation. We denote by $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ and $\mathbb{N} = \{1, 2, \dots\}$ the sets of non-negative and positive integers, respectively. We write $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$.

2. THE DERRIDA-GRIFFITHS-HIGGS DIRECTED MODEL

We recall the definition of the model in Section 2.1 and review previous results as well as conjectures made in [15] in Section 2.2. Our results are contained in Sections 2.3 to 2.5. In Section 2.3, we first show that the (infinite volume) quenched free energy exists and then provide a variational formula for it using the so-called Large Deviation Principle (LDP) *for words drawn in a letter sequence* [3, 4, 5]. Finally, we obtain some bounds on the free energy at high and low temperatures, in Sections 2.4 and 2.5 respectively.

Before we continue, let us distinguish the old material from the new one in the present paper. Proposition 2.3, Proposition 2.4 and Lemma 2.5 are taken from [15], but we take

*The word *directed* can be misleading. In the context of polymer models, it usually refers to the use of a $(1+d)$ -dimensional random walk, that is a walk which moves deterministically in the first dimension and according to simple random walk in the remaining d dimensions. It has a different meaning in this paper.

the occasion to write full proofs, obtaining thereby a (marginally) improved lower bound in Proposition 2.4. The idea behind Proposition 2.7 goes back to [15], but we obtain a more precise statement by writing a detailed proof. We have not seen the results in Theorem 2.10 and Proposition 2.11 displayed elsewhere, although they follow from standard arguments. Also, the argument behind the lower bound in Item (2) of Proposition 2.11 appeared in [15]. The other results in this section are new, to the best of the author's knowledge.

2.1. Definition of the model. We introduce two (slightly different but equivalent) conventions for the model. The convention in Section 2.1.1 corresponds to that in [15] whereas that in Section 2.1.2 looks closer to the convention used in [2, 8]. Finally, we observe that the model fits into a class of (disordered) statistical-mechanics systems built on renewal sequences.

2.1.1. Original convention. We assume that the charges $(\omega_i)_{i \geq 1}$ are i.i.d. and follow the unbiased charge distribution $\mathbb{P}(\omega_1 = 1) = \mathbb{P}(\omega_1 = -1) = 1/2$, unless stated otherwise. When summing consecutive charges we shall adopt the notation

$$(2.1) \quad \Omega_{(a,b]} = \sum_{a < i \leq b} \omega_i, \quad 0 \leq a < b,$$

and for conciseness we write Ω_n instead of $\Omega_{(0,n]}$. The set of allowed paths (i.e. the polymer configurations) is given by

$$(2.2) \quad \mathcal{S}_n = \{(s_i)_{i=0}^n : s_0 = 0, s_1 = 1, s_{i+1} - s_i \in \{0, 1\}, 1 \leq i < n\}, \quad n \geq 1,$$

and $\mathcal{S}_0 := \{0\}$. The fact that the first step is fixed to one in the above definition is only a matter of convention. In the polymer interpretation, the i -th monomer in the chain has position s_i and bears a charge ω_i . The Hamiltonian is then defined as

$$(2.3) \quad H_n^\omega(s) = \sum_{1 \leq i < j \leq n} \omega_i \omega_j \mathbf{1}_{\{s_i = s_j\}}, \quad s \in \mathcal{S}_n,$$

so that the quenched partition function writes:

$$(2.4) \quad Z_n^{\beta,\omega} = \sum_{s \in \mathcal{S}_n} \exp(-\beta H_n^\omega(s)),$$

where $\beta \geq 0$ is the *inverse temperature*. For $n \in \{0, 1\}$ the sum in (2.3) is empty, hence $H_0^\omega = H_1^\omega = 0$ and $Z_0^{\beta,\omega} = Z_1^{\beta,\omega} = 1$. We shall write $Z_n^{\beta,\omega}(A)$ when the sum in (2.4) is restricted to $s \in A \subseteq \mathcal{S}_n$. We denote by $\mathbb{P}_n^{\beta,\omega}$ the quenched polymer measure, that is

$$(2.5) \quad \mathbb{P}_n^{\beta,\omega}((S_i)_{i=0}^n = s) = \frac{\exp(-\beta H_n^\omega(s))}{Z_n^{\beta,\omega}}, \quad s \in \mathcal{S}_n,$$

and by $\mathbb{E}_n^{\beta,\omega}$ (resp. $\mathbb{V}_n^{\beta,\omega}$) the corresponding expectation (resp. variance). In the sequel we shall also use the notation

$$(2.6) \quad Z_{(a,b]}^{\beta,\omega} = Z_{b-a}^{\beta,\theta^a \omega}, \quad 0 \leq a \leq b,$$

where θ is the shift operator, i.e. $(\theta\omega)_i = \omega_{i+1}$.

2.1.2. *Alternative convention.* Let $(S_i)_{i \geq 1}$ be a directed random walk on the set of natural integers such that $S_0 = 0$, $S_1 = 1$ and $(\Delta S_i)_{i \geq 1} := (S_{i+1} - S_i)_{i \geq 1}$ is a sequence of i.i.d. random variables uniformly distributed on $\{0, 1\}$ (corresponding to *folded* or *stretched* monomers). Define the quenched partition function as

$$(2.7) \quad \bar{Z}_n^{\beta, \omega} = \mathbb{E} \left[\exp \left(-\beta \sum_{1 \leq i, j \leq n} \omega_i \omega_j \mathbf{1}_{\{S_i = S_j\}} \right) \right], \quad n \geq 1,$$

where \mathbb{E} is the expectation with respect to the law \mathbb{P} of the random walk S . It will be convenient in the sequel to set the convention $\bar{Z}_0^{\beta, \omega} = 2$. It was already noted that this partition function can be rewritten

$$(2.8) \quad \bar{Z}_n^{\beta, \omega} = \mathbb{E} \left[\exp \left(-\beta \sum_{x \geq 1} \Omega_n(x)^2 \right) \right], \quad n \in \mathbb{N},$$

where

$$(2.9) \quad \Omega_n(x) = \sum_{1 \leq i \leq n} \omega_i \mathbf{1}_{\{S_i = x\}}$$

is the cumulated charge at site $x \in \mathbb{N}$. The two conventions are actually equivalent. Indeed, the equality

$$(2.10) \quad \sum_{1 \leq i, j \leq n} \omega_i \omega_j \mathbf{1}_{\{S_i = S_j\}} = n + 2H_n^\omega$$

(that holds for *binary* charges) implies that for every $n \in \mathbb{N}_0$,

$$(2.11) \quad \bar{Z}_n^{\beta, \omega} = 2^{1-n} e^{-\beta n} Z_n^{2\beta, \omega}.$$

Consequently, the polymer measures coincide up to a different scaling of the inverse temperature:

$$(2.12) \quad \bar{\mathbb{P}}_n^{\beta, \omega} := \frac{1}{\bar{Z}_n^{\beta, \omega}} \mathbb{E} \left[\exp \left(-\beta \sum_{x \geq 1} \Omega_n(x)^2 \right) \mathbf{1}_{\{\cdot\}} \right] = \mathbb{P}_n^{2\beta, \omega},$$

and the corresponding versions of the quenched free energy are easily connected one to another:

$$(2.13) \quad \bar{F}_{\text{que}}(\beta) = -\beta - \log 2 + F_{\text{que}}(2\beta),$$

where $F_{\text{que}}(\beta) := \lim(1/n) \log Z_n^{\beta, \omega}$ and $\bar{F}_{\text{que}}(\beta) := \lim(1/n) \log \bar{Z}_n^{\beta, \omega}$ as $n \rightarrow \infty$, provided that these limits exist. We will come back to the issue of existence later in the paper and prove that such limits actually exist in the \mathbb{P} -a.s. and in the $L^1(\mathbb{P})$ sense, see Theorem 2.10 below.

2.1.3. *Renewal times.* We set $\tau_0 := 0$ and $\tau_i = \inf\{k > \tau_{i-1} : \Delta S_k = 1\}$ for every $i \geq 1$, or in other words, $\tau = \{\tau_i\}_{i \geq 0} = \{k \geq 0 : \Delta S_k = 1\}$ (we used here that $S_0 := 0$ and $S_1 := 1$). Note that (i) τ is a renewal process with $\text{Geo}(1/2)$ inter-arrival distribution on \mathbb{N} and (ii) the (quenched) partition function can be expressed in terms of this renewal process directly, instead of the random walk:

$$(2.14) \quad \bar{Z}_n^{\beta, \omega} = \mathbb{E} \left[\exp \left(-\beta \sum_{i \geq 1} \Omega(\tau_{i-1} \wedge n, \tau_i \wedge n)^2 \right) \right],$$

with the convention that $\Omega(\emptyset) = 0$. In this way we see that this directed model fits into a broader class of (disordered) statistical mechanics model that are built on renewal sequences and that have witnessed remarkable progress quite recently, such as the random

pinning model and the copolymer model, see e.g. [13, 17, 18] and references therein. We shall come back to this observation in Section 2.3.

2.2. Predictions and state of the art. In this section we recall a series of predictions and observations originally made in [15]. To this purpose, let us define

$$(2.15) \quad p_{i,n}^{\beta,\omega} := \mathbb{E}_n^{\beta,\omega}(\Delta S_i), \quad \Delta S_i = S_{i+1} - S_i \in \{0, 1\}, \quad 1 \leq i < n,$$

as well as the quenched *empirical cumulative distribution function*

$$(2.16) \quad \mathcal{F}_n^{\beta,\omega}(p) = \frac{1}{n-1} \sum_{1 \leq i < n} \mathbb{1}\{p_{i,n}^{\beta,\omega} \leq p\}, \quad p \in [0, 1], \quad n \geq 2.$$

In [15], Derrida, Griffiths and Higgs (DGH) noticed, based on numerical simulations, that “many p_i ’s have a rapid and nonmonotonic temperature variation, not unlike the “chaotic” behaviour of the local magnetization m_i in a spin glass”. More precisely, they formulated the following:

Conjecture 2.1. (*Self-averaging and “weak freezing transition”*)

- (1) *The specific heat (i.e. the derivative of the finite volume free energy) converges \mathbb{P} -almost-surely, as $n \rightarrow \infty$, to a non-random limit.*
- (2) *The empirical cumulative distribution function $\mathcal{F}_n^{\beta,\omega}(p)$ converges \mathbb{P} -almost-surely, as $n \rightarrow \infty$, to a non-random limit $\mathcal{F}_\beta(p)$.*
- (3) *There exists a critical inverse temperature $\beta_c \in (0, \infty)$ such that:*
 - *For every $\beta < \beta_c$, there exists $p_{\min}(\beta) \in (0, 1)$ such that $\mathcal{F}_\beta(p) = 0$ for every $p < p_{\min}(\beta)$ (high-temperature regime).*
 - *For every $\beta > \beta_c$, there exists an exponent $\gamma(\beta) > 0$ such that $\mathcal{F}_\beta(p) \sim (\text{cst.}) p^{\gamma(\beta)}$ as $p \rightarrow 0$ (low-temperature regime).*

Remark 2.2. *The authors in [15, Eq. (14)] provide heuristic arguments showing that the exponent $\gamma(\beta)$ continuously varies in β and solves*

$$(2.17) \quad \sum_{n \geq 1} 2^{n/\gamma} e^{-\frac{1}{2}\beta n^2} = 1.$$

In [15], the authors established:

- the presence of a *high-temperature regime* for $\beta < \log(\frac{1+\sqrt{5}}{2})$ and general charge sequences, see Section 2.2.2;
- the presence of a *low-temperature regime* for a particular di-block n -dependent charge sequence, namely (for a system of size $2n$)

$$(2.18) \quad \omega_i^{(2n)} = \begin{cases} 1 & \text{if } 1 \leq i < n, \\ -1 & \text{if } n \leq i < 2n, \end{cases}$$

see Section 2.2.3.

The remaining points in Conjecture 2.1, including the presence of a low-temperature phase for the binary (or other) charge distribution, are still open, to the best of our knowledge.

In the rest of the section, we provide full proofs for some of the results stated in [15].

2.2.1. *Recursion relation.* Recall the definition of τ in Section 2.1.3. By decomposing a random walk path according to $\max\{0 \leq k < n : k \in \tau\}$, that is the last renewal point to be found (strictly) before monomer $n \geq 1$, we readily get the following recursive relation:

$$(2.19) \quad Z_n^{\beta,\omega} = \sum_{0 \leq k < n} Z_k^{\beta,\omega} \exp\left(\frac{\beta}{2}[n - k - \Omega_{(k,n)}^2]\right),$$

or equivalently,

$$(2.20) \quad \bar{Z}_n^{\beta,\omega} = \sum_{0 \leq k < n} \bar{Z}_k^{\beta,\omega} 2^{k-n} \exp(-\beta \Omega_{(k,n)}^2).$$

As noticed in [15] this relation allows for fast numerical simulation.

2.2.2. *The high temperature regime.* Recall the definition of $p_{i,n}^{\beta,\omega}$ in (2.15).

Proposition 2.3 (DGH [15]). *For every $1 \leq i < n$ and every charge sequence ω , we have*

$$(2.21) \quad p_{i,n}^{\beta,\omega} - \frac{1}{2} \sim \beta \sum_{\substack{0 < u \leq i \\ i < v \leq n}} \omega_u \omega_v 2^{u-v}, \quad \beta \rightarrow 0.$$

Proof of Proposition 2.3. Let $1 \leq i < n$ and consider the ratio:

$$(2.22) \quad \frac{P_n^{\beta,\omega}(\Delta S_i = 1)}{P_n^{\beta,\omega}(\Delta S_i = 0)} = \frac{Z_n^{\beta,\omega}(\Delta S_i = 1)}{Z_n^{\beta,\omega}(\Delta S_i = 0)}.$$

Let us first deal with the numerator. We first note that:

$$(2.23) \quad Z_n^{\beta,\omega}(\Delta S_i = 1) = Z_i^{\beta,\omega} Z_{(i,n)}^{\beta,\omega}.$$

By decomposing a path according to $k := \max\{[0, i] \cap \tau\}$ and $\ell := \min\{(i, n) \cap \tau\}$ (by convention we set $\ell := n$ if the latter set is empty), we obtain

$$(2.24) \quad Z_n^{\beta,\omega}(\Delta S_i = 1) = \sum_{\substack{0 \leq k < i \\ i < \ell \leq n}} Z_k^{\beta,\omega} e^{\frac{\beta}{2}(\ell - k - \Omega_{(k,i)}^2 - \Omega_{(i,\ell)}^2)} Z_{(\ell,n)}^{\beta,\omega},$$

with the convention that the rightmost partition function in the line above is to be understood as equal to one if $\ell = n$. As for the denominator, a similar decomposition gives:

$$(2.25) \quad Z_n^{\beta,\omega}(\Delta S_i = 0) = \sum_{\substack{0 \leq k < i \\ i < \ell \leq n}} Z_k^{\beta,\omega} e^{\frac{\beta}{2}(\ell - k - \Omega_{(k,\ell)}^2)} Z_{(\ell,n)}^{\beta,\omega}.$$

Comparing both expressions, we observe that

$$(2.26) \quad \Omega_{(k,\ell)}^2 - \Omega_{(k,i)}^2 - \Omega_{(i,\ell)}^2 = 2 \sum_{\substack{k < u \leq i \\ i < v \leq \ell}} \omega_u \omega_v.$$

Performing a small- β expansion, and noting that $Z_n^{\beta,\omega}$ converges to $2^{(n-1) \vee 0}$ as $\beta \rightarrow 0$, we obtain (the case $k = 0$ or $\ell = n$ needs care):

$$(2.27) \quad p_{i,n}^{\beta,\omega} - \frac{1}{2} \sim \frac{\beta}{2} \sum_{\substack{0 \leq k < i \\ i < \ell \leq n}} 2^{k \vee 1 - \ell \wedge (n-1)} \sum_{\substack{k < u \leq i \\ i < v \leq \ell}} \omega_u \omega_v, \quad \beta \rightarrow 0.$$

We conclude by noting that

$$(2.28) \quad \sum_{\substack{0 \leq k < i \\ i < \ell \leq n}} 2^{k \vee 1 - \ell \wedge (n-1)} \sum_{\substack{k < u \leq i \\ i < v \leq \ell}} \omega_u \omega_v = \sum_{\substack{0 < u \leq i \\ i < v \leq n}} \omega_u \omega_v 2^{1+u-v}.$$

■

Let us make a few comments on Proposition 2.3. First, we observe that the series in the r.h.s. of (2.21) converges, under mild assumptions on the charge sequence (say boundedness), as $n \rightarrow \infty$:

$$(2.29) \quad \lim_{n \rightarrow \infty} \sum_{\substack{0 < u \leq i \\ i < v \leq n}} \omega_u \omega_v 2^{u-v} = \sum_{0 < u \leq i < v} \omega_u \omega_v 2^{u-v}.$$

This shows that far-away charges have a decreasing influence on the state of a given monomer, at high temperatures. However, it is not clear whether the limits $\beta \rightarrow 0$ and $n \rightarrow \infty$ may be interchanged. Using the same proof as above but replacing the small- β expansion by the inequality $e^x \geq 1 + x$, one can interchange the limits and obtain:

$$(2.30) \quad \limsup_{\beta \rightarrow 0} \frac{1}{\beta} \left(p_{i,n}^{\beta,\omega} - \frac{1}{2} \right) \leq \sum_{\substack{0 < u \leq i \\ i < v \leq n}} \omega_u \omega_v 2^{u-v},$$

uniformly in $1 \leq i < n$. This is however not the desired direction, in view of Item (3) in Conjecture 2.1. Fortunately, DGH noticed the following:

Proposition 2.4 (DGH [15]). *For every $1 \leq i < n$ and every charge sequence ω ,*

$$(2.31) \quad p_{i,n}^{\beta,\omega} \geq \left(1 + \frac{e^{\beta/2} + e^{\beta}[1 + e^{-\beta}]^{-1}}{1 - e^{\beta}[1 + e^{-\beta}]^{-1}} \right)^{-1},$$

which is positive for every $\beta < \log(\frac{1+\sqrt{5}}{2})$.

The proof relies on the following lemma:

Lemma 2.5 (DGH [15]). *For every $n \geq 1$, $Z_n^{\beta,\omega} \geq Z_{n-1}^{\beta,\omega}$ and for every $n \geq 2$, $Z_n^{\beta,\omega} \geq (1 + e^{-\beta}) Z_{n-2}^{\beta,\omega}$.*

Let us stress that Lemma 2.5 holds for any charge sequence ω .

Proof of Lemma 2.5. Recall the expression of the partition function in (2.4). The proof follows by restricting \mathcal{S}_n to $\{s \in \mathcal{S}_n : s_n - s_{n-1} = 1\}$ for the first inequality, and $\{s \in \mathcal{S}_n : s_{n-1} - s_{n-2} = 1\}$ for the second one. ■

Proof of Proposition 2.4. Starting from (2.25) and ignoring the term $\Omega_{(k,\ell)}^2$ in the exponential, we get:

$$(2.32) \quad Z_n^{\beta,\omega}(\Delta S_i = 0) \leq \sum_{0 \leq k < i} Z_k^{\beta,\omega} e^{\frac{\beta}{2}(i-k)} \sum_{i < \ell \leq n} e^{\frac{\beta}{2}(\ell-i)} Z_{(\ell,n)}^{\beta,\omega}.$$

Using Lemma 2.5 repeatedly, we have for every $0 \leq u \leq (n-1)/2$,

$$(2.33) \quad \max(Z_{n-2u}^{\beta,\omega}, Z_{n-2u-1}^{\beta,\omega}) \leq (1 + e^{-\beta})^{-u} Z_n^{\beta,\omega}.$$

Plugging this estimate into the sum over k and splitting it between even and odd values of $i - k$ (say $i - k = 2u + 1$ for some $u \geq 0$ or $i - k = 2u$ for some $u \geq 1$) we obtain:

$$(2.34) \quad \sum_{0 \leq k < i} Z_k^{\beta, \omega} e^{\frac{\beta}{2}(i-k)} \leq Z_i^{\beta, \omega} \left[\sum_{u \geq 0} (1 + e^{-\beta})^{-u} e^{(2u+1)\beta/2} + \sum_{u \geq 1} (1 + e^{-\beta})^{-u} e^{(2u)\beta/2} \right] \\ = Z_i^{\beta, \omega} \left[\frac{e^{\beta/2} + e^{\beta}[1 + e^{-\beta}]^{-1}}{1 - e^{\beta}[1 + e^{-\beta}]^{-1}} \right],$$

as soon as $\beta < \log(\frac{1+\sqrt{5}}{2})$, so that the sum over u is finite. Similarly,

$$(2.35) \quad \sum_{i < \ell \leq n} e^{\frac{\beta}{2}(\ell-i)} Z_{(\ell, n)}^{\beta, \omega} \leq Z_{(i, n)}^{\beta, \omega} \left[\frac{e^{\beta/2} + e^{\beta}[1 + e^{-\beta}]^{-1}}{1 - e^{\beta}[1 + e^{-\beta}]^{-1}} \right].$$

Combining (2.23), (2.32), (2.34) and (2.35) leads to

$$(2.36) \quad Z_n^{\beta, \omega}(\Delta S_i = 0) \leq \left[\frac{e^{\beta/2} + e^{\beta}[1 + e^{-\beta}]^{-1}}{1 - e^{\beta}[1 + e^{-\beta}]^{-1}} \right]^2 \times Z_n^{\beta, \omega}(\Delta S_i = 1),$$

which completes the proof. ■

Remark 2.6. *Although we follow the same line of proof, the lower bound in Proposition 2.4 is (slightly) better than the one stated in [15]. However, none of these bounds is close to optimal, as the limit $\beta \rightarrow 0$ shows.*

2.2.3. The low temperature regime. Although the existence of a low-temperature regime remained at a heuristic level for the random binary charges, DGH [15] shortly argued that for the special charge distribution defined in (2.18) (a.k.a. di-block polymer), the chain collapses at inverse temperatures $\beta > 2\beta_0$, where

$$(2.37) \quad \beta_0 := \inf \left\{ \beta > 0 : S(\beta) := \sum_{t \in \mathbb{N}} e^{-\beta t^2} < 1 \right\},$$

meaning that “almost all of the monomers are on one site, as the cancellation of positive and negative charges minimizes the energy”. In this section, we provide the necessary details of this argument so as to obtain a more quantitative statement. To this purpose, let us define

$$(2.38) \quad i(n) = \max([0, n] \cap \tau) \\ j(n) = \min([n, 2n] \cap \tau) - n,$$

with the convention $j(n) = n$ if $[n, 2n] \cap \tau = \emptyset$.

Proposition 2.7. *Assume that the charge sequence is as in (2.18) and that $\beta > \beta_0$. Then, there exists $C(\beta) \in (0, 1)$ such that for every $0 \leq i < n$ and $0 \leq j \leq n$,*

$$(2.39) \quad \bar{P}_{2n}^{\beta, \omega}(i(n) = i) \leq \frac{C(\beta)^i}{1 - S(\beta)}, \quad \bar{P}_{2n}^{\beta, \omega}(j(n) = j) \leq \frac{C(\beta)^{n-j}}{1 - S(\beta)}$$

and for every $M \geq 1$,

$$(2.40) \quad \bar{P}_{2n}^{\beta, \omega}(|i(n) + j(n) - n| \geq M) \leq \frac{e^{-\beta M^2}}{[1 - S(\beta)]^2}.$$

Corollary 2.8. *Under the same assumptions as Proposition 2.7, we have*

$$(2.41) \quad P_{i, 2n}^{2\beta, \omega} \leq \frac{C(\beta)^{i \wedge (2n-i)}}{[1 - S(\beta)][1 - C(\beta)]}, \quad 1 \leq i \leq 2n.$$

By (2.12), the threshold value β_0 in Proposition 2.7 indeed corresponds to $2\beta_0$ in [15]. We will see during the proof that the rate of decay $C(\beta)$ can be expressed in terms of the free energy of the *weakly self-avoiding walk* (defined slightly below). Proposition 2.7 shows the presence of a low-temperature regime (for this particular charge sequence) where (i) the polymer is very close to its *fully collapsed* state (corresponding to $i(n) = 0$ and $j(n) = n$) and (ii) the cumulated charge in the folded piece of polymer around monomer n (equal to $i(n) + j(n) - n$) has a Gaussian tail. Before starting the proof, let us introduce some additional notation. We write

$$(2.42) \quad \bar{Z}_n^{\beta,+} := \mathbb{E} \left[\exp \left(-\beta \sum_{1 \leq i, j \leq n} \mathbf{1}_{\{S_i = S_j\}} \right) \right], \quad n \geq 1,$$

that is the partition function displayed in (2.7) when all the charges equal one (a.k.a. the partition function of the *weakly self-avoiding walk*). It can be rewritten as

$$(2.43) \quad \bar{Z}_n^{\beta,+} = \mathbb{E} \left[\exp \left(-\beta \sum_{x \geq 1} \ell_n(x)^2 \right) \right], \quad \text{with } \ell_n(x) := \{1 \leq k \leq n : S_k = x\},$$

with the convention $\bar{Z}_0^{\beta,+} = 2$, in order to be consistent with Section 2.1.2. We shall use the following lemma, the proof of which is deferred at the end of this section. Recall the definition of β_0 in (2.37).

Lemma 2.9. *The following limit exists*

$$(2.44) \quad -\bar{F}_+(\beta) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \bar{Z}_n^{\beta,+} = \sup_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{1}{2} \bar{Z}_n^{\beta,+} \right),$$

and $\bar{F}_+(\beta) > \log 2$ if and only if $\beta > \beta_0$.

Proof of Proposition 2.7. We first observe that for every $0 \leq i < n$ and $0 \leq j \leq n$,

$$(2.45) \quad \bar{Z}_{2n}^{\beta,\omega}(i(n) = i, j(n) = j) = \bar{Z}_i^{\beta,+} \times [2^{i-j-n-1} e^{-\beta(i+j-n)^2}] \times Z_{n-j}^{\beta,+},$$

by the Markov property at times i and j , and “reversing time” on the interval $(j, n]$. The *fully collapsed* state, corresponding to $i(n) = 0$ and $j(n) = n$, yields the following lower bound:

$$(2.46) \quad \bar{Z}_{2n}^{\beta,\omega} \geq \bar{Z}_{2n}^{\beta,\omega}(i(n) = 0, j(n) = n) = 2^{1-2n}.$$

(We remind the reader that the first step of the random walk is fixed to the value one). Therefore,

$$(2.47) \quad \bar{P}_{2n}^{\beta,\omega}(i(n) = i, j(n) = j) \leq [2^{i-1} \bar{Z}_i^{\beta,+}] e^{-\beta(i+j-n)^2} [2^{n-j-1} \bar{Z}_{n-j}^{\beta,+}].$$

By Lemma 2.9, we have, since $\beta > \beta_0$,

$$(2.48) \quad 2^{k-1} \bar{Z}_k^{\beta,+} \leq C(\beta)^k, \quad \text{with } C(\beta) := 2e^{-\bar{F}_+(\beta)} \in (0, 1),$$

and (using (2.52) below and the definition of $S(\beta)$ in (2.37))

$$(2.49) \quad \sum_{k \geq 0} 2^{k-1} \bar{Z}_k^{\beta,+} = \frac{1}{1 - S(\beta)},$$

which completes the proof. ■

Proof of Lemma 2.9. The first part of the lemma is rather standard and follows from Fekete’s lemma, once we notice that the sequence defined by

$$(2.50) \quad \bar{Z}_n^{\beta,+}(n \in \tau) = \bar{Z}_n^{\beta,+}(\Delta S_n = 1) = (1/2) \bar{Z}_n^{\beta,+}, \quad n \geq 1,$$

is super-multiplicative. The second part of the lemma follows from

$$(2.51) \quad \begin{aligned} \bar{Z}_n^{\beta,+}(n \in \tau) &= \mathbb{E} \left[\exp \left(-\beta \sum_{x \geq 1} \ell_n(x)^2 \right) \mathbf{1}_{\{n \in \tau\}} \right] \\ &= \sum_{k \geq 1} \sum_{0=t_0 < t_1 < \dots < t_k=n} \prod_{1 \leq i \leq k} e^{-\beta(t_i - t_{i-1})^2} 2^{-(t_i - t_{i-1})} \end{aligned}$$

which implies that

$$(2.52) \quad \sum_{n \geq 1} e^{fn} \bar{Z}_n^{\beta,+}(n \in \tau) = \sum_{k \geq 1} \left(\sum_{t \geq 1} e^{(f - \log 2)t - \beta t^2} \right)^k, \quad f \in \mathbb{R}.$$

■

2.3. Existence and variational representation of the free energy. In this section we focus on the quenched free energy of the model, using one or the other definition, thanks to (2.13). We start with the issue of existence and self-averaging in Section 2.3.1, along with a few preliminary properties of the free energy as a function of the inverse temperature. In Section 2.3.2 we apply the so-called *Large Deviation Principle (LDP) for words drawn in a random letter sequence* [3, 4, 5] to derive a variational representation of the free energy. We will use this representation in Section 2.4 and obtain new estimates on the free energy.

2.3.1. Existence and self-averaging. Similarly to how we proved Lemma 2.9, we define a *constrained* version of the quenched partition function by:

$$(2.53) \quad \bar{Z}_{n,c}^{\beta,\omega} := \bar{Z}_n^{\beta,\omega}(n \in \tau) = \bar{Z}_n^{\beta,\omega}(\Delta S_n = 1).$$

Since the value of ΔS_n does not affect the Hamiltonian, we readily get

$$(2.54) \quad \bar{Z}_{n,c}^{\beta,\omega} = \frac{1}{2} \bar{Z}_n^{\beta,\omega}.$$

We prove the following:

Theorem 2.10. *The quenched free energy exists as the non-random limit*

$$(2.55) \quad \bar{F}_{\text{que}}(\beta) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \bar{Z}_n^{\beta,\omega} = \lim_{n \rightarrow \infty} \frac{1}{n} \log \bar{Z}_{n,c}^{\beta,\omega} \leq 0,$$

which holds \mathbb{P} -almost surely and in $L^1(\mathbb{P})$. Moreover,

$$(2.56) \quad \bar{F}_{\text{que}}(\beta) = \sup_{n \geq 1} \frac{1}{n} \mathbb{E} \log \bar{Z}_{n,c}^{\beta,\omega}.$$

Proof of Theorem 2.10. Notice that $\bar{Z}_{n+m,c}^{\beta,\omega} \geq \bar{Z}_{n,c}^{\beta,\omega} \bar{Z}_{m,c}^{\beta,\omega}$, apply the logarithm function, and conclude via Kingman's superadditive ergodic theorem. ■

Let us list some properties of the quenched free energy as a function of the inverse temperature (see Figures 1 and 2):

Proposition 2.11. *We have the following:*

- (1) $\beta \in [0, \infty) \mapsto \bar{F}_{\text{que}}(\beta)$ is a convex non-increasing function.
- (2) $\bar{F}_{\text{que}}(\beta) \geq (-\beta) \vee (-\log 2)$.

We do not know whether the quenched free energy is differentiable as a function of the inverse temperature. If it were so, then by convexity the derivative would be the limit of the derivative of the finite-volume free energy, as $n \rightarrow \infty$ (see Item (1) in Conjecture 2.1).

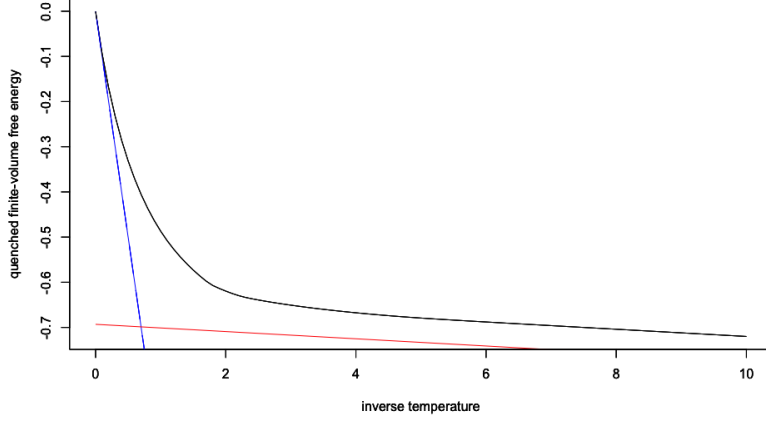


FIGURE 1. Quenched finite-volume free energy $(1/n) \log \bar{Z}_n^{\beta, \omega}$ versus inverse temperature β for centered ± 1 i.i.d. charges ω and polymer size $n = 1000$ as defined in (2.8). The high-temperature lower bound $(-\beta)$ and the low-temperature lower bound from (2.61) respectively appear in blue and red.

Proof of Proposition 2.11. For (1) it is enough to check that

$$(2.57) \quad \begin{aligned} \partial_\beta \left(\frac{1}{n} \log \bar{Z}_n^{\beta, \omega} \right) &= -\bar{\mathbb{E}}_n^{\beta, \omega} \left(\sum_{x \geq 1} \Omega_n(x)^2 \right) \leq 0, \\ \partial_\beta^2 \left(\frac{1}{n} \log \bar{Z}_n^{\beta, \omega} \right) &= \bar{\mathbb{V}}_n^{\beta, \omega} \left(\sum_{x \geq 1} \Omega_n(x)^2 \right) \geq 0. \end{aligned}$$

Let us deal with (2). By Jensen's inequality and the i.i.d. assumption on ω ,

$$(2.58) \quad \mathbb{E} \log \bar{Z}_n^{\beta, \omega} \geq -\beta \mathbb{E} \mathbb{E} \left(\sum_{1 \leq i, j \leq n} \omega_i \omega_j \mathbf{1}_{\{S_i = S_j\}} \right) = -\beta n,$$

which yields the first inequality. To obtain the second inequality, let us first assume for simplicity that $\Omega_n > 0$ and define, for every $1 \leq i \leq \Omega_n$,

$$(2.59) \quad \sigma_i = \sup\{1 \leq k \leq n : \Omega_k = i\}.$$

By considering the one particular path which satisfies $\Delta S_k = 1$ if and only if $k \in \sigma = \{\sigma_i, 1 \leq i \leq \Omega_n\}$, for $1 \leq k \leq n$ (i.e. $\tau_i = \sigma_i$ for $1 \leq i \leq \Omega_n$), we obtain

$$(2.60) \quad \frac{1}{n} \log \bar{Z}_n^{\beta, \omega} \geq -\log 2 - \beta \frac{\Omega_n}{n}.$$

One may readily adapt the argument to the case $\Omega_n \leq 0$ and obtain

$$(2.61) \quad \frac{1}{n} \log \bar{Z}_n^{\beta, \omega} \geq -\log 2 - \beta \frac{|\Omega_n|}{n},$$

leading to the following lower bound:

$$(2.62) \quad \frac{1}{n} \mathbb{E} \log \bar{Z}_n^{\beta, \omega} \geq -\log 2 - \beta \frac{\mathbb{E}|\Omega_n|}{n} = -\log 2 + o(1),$$

which completes the proof of (2). ■

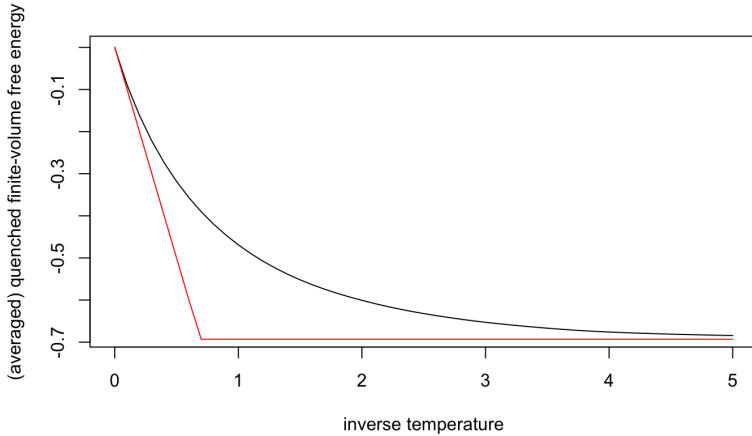


FIGURE 2. Averaged quenched finite-volume free energy $(1/n)\mathbb{E} \log \bar{Z}_n^{\beta, \omega}$ versus inverse temperature β for centered ± 1 i.i.d. charges ω and polymer size $n = 1000$ as defined in (2.8). We have used 100 samples to approximate the average over the charge distribution. The lower bound stated in Proposition 2.11 is shown in red.

2.3.2. *The LDP for words in a quenched random letter sequence.* As we already observed in Section 2.1.3, the present model can be formulated as a statistical mechanics model built on a renewal sequence in the presence of quenched randomness, like the random pinning and copolymer models [13, 17, 18]. In these models, one can visualize the random charge sequence $(\omega_i)_{i \geq 1}$ as a random *letter sequence* from which the renewal process $(\tau_k)_{k \geq 0}$ cuts a *word sequence* $(Y_k)_{k \geq 1}$, where $Y_k := (\omega_i)_{\tau_{k-1} < i \leq \tau_k}$ for every $k \geq 1$. Then, the energy or Hamiltonian function may be recast as an appropriate functional applied to the *word sequence empirical measure*. The purpose of this section is to apply a large deviation principle (LDP) for the empirical measure of such words (when the letter sequence is quenched) in order to obtain a variational formula for the free energy. This LDP at the level of measures was originally introduced by Birkner [3], shortly later generalized to a broader class of renewal tails by Birkner, Greven, den Hollander [4, 5] and successfully applied to the random pinning and copolymer models [6, 7, 10, 14, 20].

To implement this idea, let us first define the grand canonical partition function:

$$(2.63) \quad \mathcal{Z}_{\beta, \omega}(f) := \sum_{n \geq 1} \bar{Z}_{n, c}^{\beta, \omega} e^{-fn}, \quad f \in \mathbb{R}.$$

Note that we use the constrained version of the partition function defined in (2.53). By Theorem 2.10, the free energy is the infimum of those values of f which make the latter sum converge:

$$(2.64) \quad \bar{F}_{\text{que}}(\beta) = \inf \{f \in \mathbb{R} : \mathcal{Z}_{\beta, \omega}(f) < \infty\}.$$

By (2.14), we may write

$$\begin{aligned}
 \mathcal{Z}_{\beta,\omega}(f) &= \sum_{n \geq 1} e^{-fn} \sum_{k \geq 1} \mathbb{E} \left[\exp \left(-\beta \sum_{1 \leq i \leq k} \Omega(\tau_{i-1}, \tau_i)^2 \right) \mathbf{1}_{\{\tau_k = n\}} \right] \\
 (2.65) \qquad &= \sum_{k \geq 1} \mathbb{E} \left[\exp \left(-\sum_{i=1}^k [\beta \Omega(\tau_{i-1}, \tau_i)^2 + f(\tau_i - \tau_{i-1})] \right) \right] =: \sum_{k \geq 1} W_k^{\beta,\omega}.
 \end{aligned}$$

Using the Large Deviation Principle for words cut out of a quenched random sequence of letters and Varadhan's lemma [3, Theorem 1 and Corollary 1] we obtain the following \mathbb{P} -a.s. limit:

$$(2.66) \qquad \lim_{k \rightarrow \infty} \frac{1}{k} \log W_k^{\beta,\omega} = -J(\beta, f), \qquad f > -\log 2,$$

where

$$(2.67) \qquad J(\beta, f) := \inf_{Q \in \mathcal{R}_0} \int [\beta(y_{1,1} + \dots + y_{1,\ell(y_1)})^2 + f\ell(y_1)] Q(dy) + H(Q|Q_0)$$

and with the following notation:

- Q is a probability distribution on sequences of *words*. Words are elements of $\cup_{\ell \geq 1} \mathbb{R}^\ell$, and $\ell(y_1)$ denotes the length of $y_1 = (y_{1,1}, \dots, y_{1,\ell(y_1)})$, that is the first word in the sequence $y = (y_i)_{i \geq 1}$.
- Q_0 is the probability distribution for i.i.d. words with marginal

$$(2.68) \qquad Q_0(dy_1) = \sum_{\ell \geq 1} 2^{-\ell} \prod_{1 \leq i \leq \ell} \mathbb{P}(dy_{1,i}),$$

that is the one induced by the charge/letter sequence law \mathbb{P} (Section 2.1.1) and the geometrically distributed inter-arrival times $(\tau_k - \tau_{k-1})_{k \geq 1}$ (Section 2.1.3).

- $H(Q|Q_0)$ is the *specific relative entropy* of Q w.r.t. Q_0 , defined as

$$(2.69) \qquad H(Q|Q_0) = \lim_{n \rightarrow \infty} \frac{1}{n} h(Q[(Y_1, \dots, Y_n) \in \cdot] | Q_0[(Y_1, \dots, Y_n) \in \cdot]),$$

where $h(\cdot|\cdot)$ is the usual relative entropy of one probability measure w.r.t. another and the limit is non-decreasing.

- \mathcal{R}_0 denotes the set of shift-invariant distributions on word sequences which are *compatible with* \mathbb{P} , meaning that when one concatenates a typical word sequence into a letter sequence, the empirical distribution of the latter (at the level of processes) converges weakly to \mathbb{P} , see [3, Equation (8)] for a formal definition.

Note that $J(\beta, f)$ is non-decreasing in both its variables. The fact that (2.66) is not stated for all the possible values of f comes from the lack of boundedness of the function to which we apply Varadhan's lemma. As in [3] (see the remark around Eq. (19) therein) we use instead an exponential tightness property [12, Condition (4.3.3) in Theorem 4.3.1], hence the restriction $f > -\log 2$. This restriction is harmless since we know from a separate argument that $\bar{F}_{\text{que}}(\beta) \geq -\log 2$ for every $\beta > 0$, see Proposition 2.11. Combining (2.64), (2.65) and (2.66), we finally obtain the following:

Theorem 2.12. *For every $\beta > 0$,*

$$(2.70) \qquad \bar{F}_{\text{que}}(\beta) = \inf \{ f \in \mathbb{R} : J(\beta, f) > 0 \}.$$

Let us comment on a few elementary bounds obtained from this formula. By plugging $Q = Q_0$ into (2.67) we obtain that $J(\beta, f) \leq 2(\beta + f)$, from which we retrieve the elementary Jensen bound $\bar{F}_{\text{que}}(\beta) \geq -\beta$. One can also try to restrict the infimum in (2.67) to those Q 's under which words are i.i.d. with marginal of the form

$$(2.71) \quad Q(dy_1) = \sum_{\ell \geq 1} K(\ell) \prod_{1 \leq i \leq \ell} \mathbb{P}(dy_{1,i}),$$

where K ranges over all probability distributions on the set of positive integers. We obtain thereby

$$(2.72) \quad J(\beta, f) \leq \inf_K \{(f + \log 2 + \beta)m(K) - h(K)\},$$

where $m(K) := \sum_{\ell \geq 1} \ell K(\ell)$ and $h(K) := -\sum_{\ell \geq 1} K(\ell) \log K(\ell)$ are the mean and entropy of K , respectively. A simple argument (using Lagrange multipliers) shows that the optimal K follows a geometric distribution, i.e. $K(\ell) = \theta^{\ell-1}(1-\theta)$ for some $\theta \in (0, 1)$. Therefore,

$$(2.73) \quad J(\beta, f) \leq \inf_{\theta \in (0,1)} \left\{ \frac{f + \log 2 + \beta}{1 - \theta} + \left(\frac{\theta}{1 - \theta} \right) \log \theta + \log(1 - \theta) \right\}.$$

The latter is optimal at $\theta_0 := \exp(-[f + \log 2 + \beta])$, provided $\theta_0 \in (0, 1)$, leading in that case to $J(\beta, f) \leq \log(\exp(f + \log 2 + \beta) - 1)$, and ultimately to $\bar{F}_{\text{que}}(\beta) \geq -\beta$. Unfortunately, this is not any better than the Jensen lower bound. The reason is that this strategy does not take under consideration the charge sequence. In Section 2.4, we shall implement a more refined strategy where each monomer looks at the charge of the following one to decide its state (folded or stretched) and improve thereby our lower bound on the free energy.

2.4. High-temperature estimate on the free energy. This section is primarily focussed on high-temperature estimates on the quenched free energy of the charged polymer, even if some results that we derive *en route* (namely Theorem 2.13 and Proposition 2.94) are valid at all temperatures.

2.4.1. A lower bound. The starting point of our lower bound is Theorem 2.12. By testing an appropriate probability distribution on word sequences in the variational formula (2.67), we obtain an upper bound on the rate function $J(\beta, f)$, which in turn yields a lower bound on the free energy. The distribution $Q \in \mathcal{R}_0$ which we test is a tilt of the Geo(1/2) distribution which looks at pairs of consecutive charges to decide on where to fold the polymer chain, see (2.78) below. In view of Proposition 2.3, it is plausible that a better strategy would look at *all* pairs of charges, with a decreasing influence for charges that are far-apart, but the implementation of such strategy seems technically much more demanding.

Theorem 2.13. *For every $\beta > 0$,*

$$(2.74) \quad \bar{F}_{\text{que}}(\beta) \geq -\beta + \sup_{0 \leq u \leq \log 2} \left[\beta \frac{\sinh(u)}{1 + \frac{\sinh(u)}{2}} - \frac{1}{2} u \sinh(u) - \eta(u) \right],$$

where

$$(2.75) \quad \eta(u) := \frac{1}{4} [\log(2 - e^u)(2 - e^u) + \log(2 - e^{-u})(2 - e^{-u})].$$

We readily deduce thereof the following high temperature lower bound:

Corollary 2.14. *As $\beta \rightarrow 0$,*

$$(2.76) \quad \bar{F}_{\text{que}}(\beta) \geq -\beta + \frac{\beta^2}{2}[1 + o(1)].$$

Remark 2.15. *Expanding the finite-volume free energy at high temperature (see Proposition B.1 for the precise assumptions) and (blindly) interchanging the $\beta \rightarrow 0$ and $n \rightarrow \infty$ limits leads to the prediction:*

$$(2.77) \quad \bar{F}_{\text{que}}(\beta) \stackrel{(?)}{=} -\beta + \frac{4\beta^2}{3}[1 + o(1)].$$

However, such naive expansions do not always give the correct values for high-temperatures limits, see e.g. [1] in the context of pinning and copolymer models.

Proof of Theorem 2.13. For every $u \in [0, \log 2]$, we introduce a law on word sequences $Y = (Y_i)_{i \geq 1}$, denoted by Q_u , such that (i) the letter sequence distribution equals the original charge sequence distribution, i.e. $Q_u(\omega = \kappa(Y) \in \cdot) = \mathbb{P}(\omega \in \cdot)$, where κ is the *word-to-letter sequence concatenation*, and (ii) the length of the first word, denoted by L , is distributed as

$$(2.78) \quad Q_u(L = \ell | \omega) = \prod_{1 \leq i < \ell} (1 - \frac{1}{2}e^{u\omega_i\omega_{i+1}}) \frac{1}{2}e^{u\omega_\ell\omega_{\ell+1}}, \quad \ell \geq 1.$$

In other words, $L = \inf\{i \geq 1 : \theta_i = 1\}$ where $(\theta_i)_{i \geq 1}$ is a sequence of Bernoulli random variables that are independent (conditionally to ω) and such that

$$(2.79) \quad Q_u(\theta_i = 1 | \omega) = \frac{1}{2}e^{u\omega_i\omega_{i+1}}, \quad i \geq 1.$$

A similar formula is assumed to hold for the length of the next words, simply by appropriately shifting the letter sequence ω on the r.h.s. of (2.78). Note that the random variables $(\omega_i\omega_{i+1})_{i \geq 1}$ are actually independent and uniformly distributed on $\{-1, 1\}$. Following this remark and averaging (2.78) over ω yields

$$(2.80) \quad Q_u(L = \ell) = \left(1 - \frac{\cosh(u)}{2}\right)^{\ell-1} \frac{\cosh(u)}{2}, \quad \ell \geq 1,$$

that is the geometric distribution with expectation $2/\cosh(u)$. The reader may check that under Q_u the word sequence is a Markov chain with transitions

$$(2.81) \quad \mathfrak{Q}_u((\omega_i)_{1 \leq i \leq \ell_1}, (\omega_i)_{\ell_1 < i \leq \ell_2}) := 2^{\ell_1 - \ell_2} \left(\frac{1}{2}e^{u\omega_{\ell_1}\omega_{\ell_1+1}}\right) \prod_{\ell_1 < i < \ell_2} (1 - \frac{1}{2}e^{u\omega_i\omega_{i+1}})$$

(the distribution of a given word depends on the last letter of the previous word) and started from the invariant word (probability) distribution

$$(2.82) \quad \pi_u(\omega_1, \dots, \omega_\ell) := 2^{-(\ell+1)} \cosh(u) \prod_{1 \leq i < \ell} (1 - \frac{1}{2}e^{u\omega_i\omega_{i+1}}).$$

Moreover, words are i.i.d. if and only if $u = 0$, in which case one retrieves the law Q_0 defined in (2.68). From all these observations we deduce that $Q_u \in \mathcal{R}_0$, that is the restriction set of shift-invariant word sequence distributions which are compatible with \mathbb{P} , see the definition below (2.67).

We may now plug $Q = Q_u$ into (2.67), so that we obtain

$$(2.83) \quad J(\beta, f) \leq \beta \int (\omega_1 + \dots + \omega_L)^2 dQ_u + f \int L dQ_u + H(Q_u | Q_0).$$

We begin with computing the (specific) relative entropy. By (2.69) and the ergodic theorem, we have

$$(2.84) \quad \begin{aligned} H(Q_u|Q_0) &= \int \log \mathfrak{Q}_u(Y_1, Y_2) dQ_u - \int \log \pi_0(Y_1) dQ_u \\ &= \int \left[\sum_{1 \leq i < L} \log(2 - e^{u\omega_i\omega_{i+1}}) + u\omega_L\omega_{L+1} \right] dQ_u. \end{aligned}$$

With the help of Lemma 2.16 below, we find that

$$(2.85) \quad H(Q_u|Q_0) = \frac{2}{\cosh(u)} \left(\eta(u) + \frac{1}{2} u \sinh(u) \right).$$

Using once again Lemma 2.16, we may compute the contribution of this strategy to the energy term:

$$(2.86) \quad \begin{aligned} \int (\omega_1 + \dots + \omega_L)^2 dQ_u &= \int L dQ_u + 2 \sum_{1 \leq i < j} \int \omega_i \omega_j \mathbf{1}_{\{L \geq j\}} dQ_u \\ &= \frac{2}{\cosh(u)} \left(1 - \frac{\sinh(u)}{1 + \frac{\sinh(u)}{2}} \right). \end{aligned}$$

Summing up the different terms, we finally obtain

$$(2.87) \quad J(\beta, f) \leq \frac{2}{\cosh(u)} \left[\beta \left(1 - \frac{\sinh(u)}{1 + \frac{\sinh(u)}{2}} \right) + \frac{1}{2} u \sinh(u) + \eta(u) + f \right].$$

Hence, by Theorem 2.12,

$$(2.88) \quad \bar{F}_{\text{que}}(\beta) \geq \beta \left(\frac{\sinh(u)}{1 + \frac{\sinh(u)}{2}} - 1 \right) - \frac{1}{2} u \sinh(u) - \eta(u),$$

and the proof is complete. ■

Proof of Corollary 2.14. Let us pick $u = C\beta$ in Theorem 2.13, with $C > 0$ to be determined later. Noting that, as $u \rightarrow 0$,

$$(2.89) \quad \frac{\sinh(u)}{1 + \frac{\sinh(u)}{2}} \sim u, \quad u \sinh(u) \sim u^2, \quad \eta(u) = o(u^2),$$

we obtain, as $\beta \rightarrow 0$,

$$(2.90) \quad \bar{F}_{\text{que}}(\beta) \geq -\beta + \beta^2(C - C^2/2) + o(\beta^2),$$

which we optimise by setting $C = 1$. ■

Lemma 2.16. *For every $\Phi: \{-1, 1\} \mapsto \mathbb{R}$ and $u \in [0, \log 2]$,*

$$(2.91) \quad \begin{aligned} \int \Phi(\omega_i \omega_{i+1}) \mathbf{1}_{\{L > i\}} dQ_u &= \left(1 - \frac{\cosh(u)}{2} \right)^{i-1} \frac{\Phi(-1)(2 - e^{-u}) + \Phi(1)(2 - e^u)}{4}, \\ \int \Phi(\omega_i \omega_{i+1}) \mathbf{1}_{\{L = i\}} dQ_u &= \left(1 - \frac{\cosh(u)}{2} \right)^{i-1} \frac{\Phi(-1)e^{-u} + \Phi(1)e^u}{4}, \end{aligned}$$

and for every $1 \leq i < j$,

$$(2.92) \quad \int \omega_i \omega_j \mathbf{1}_{\{L \geq j\}} dQ_u = \left(1 - \frac{\cosh(u)}{2} \right)^{i-1} \left(-\frac{\sinh(u)}{2} \right)^{j-i}.$$

Proof of Lemma 2.16. Let us start with (2.91). By conditioning on ω and using (2.78),

$$(2.93) \quad \begin{aligned} \int \Phi(\omega_i \omega_{i+1}) \mathbf{1}_{\{L>i\}} dQ_u &= \int \Phi(\omega_i \omega_{i+1}) \prod_{1 \leq j \leq i} (1 - \frac{1}{2} e^{u\omega_j \omega_{j+1}}) d\mathbb{P}, \\ \int \Phi(\omega_i \omega_{i+1}) \mathbf{1}_{\{L=i\}} dQ_u &= \int \Phi(\omega_i \omega_{i+1}) \prod_{1 \leq j < i} (1 - \frac{1}{2} e^{u\omega_j \omega_{j+1}}) \frac{1}{2} e^{u\omega_i \omega_{i+1}} d\mathbb{P}. \end{aligned}$$

We conclude by observing that the random variables $(\omega_i \omega_{i+1})_{i \geq 1}$ are independent and uniformly distributed in $\{-1, 1\}$. The equality in (2.92) is proven in a similar fashion, once we notice that $\omega_i \omega_j = \prod_{i \leq k < j} \omega_k \omega_{k+1}$. \blacksquare

2.4.2. An upper bound. In this section we derive several upper bounds on the quenched free energy through *annealing*. We start with the following:

Proposition 2.17. *For every $\beta \geq 0$,*

$$(2.94) \quad \bar{F}_{\text{que}}(\beta) \leq \bar{F}_{\text{ann}}(\beta) := -\sup \left\{ f \geq 0 : \sum_{\ell \geq 1} (e^f/2)^\ell \mathbb{E}(e^{-\beta \Omega_\ell^2}) < 1 \right\} \in [-\log 2, 0].$$

Proof of Proposition 2.17. For ease of notation, let us set

$$(2.95) \quad g_\beta(\ell) := \mathbb{E}(e^{-\beta \Omega_\ell^2}) = \mathbb{E}(e^{-\beta(\omega_1 + \dots + \omega_\ell)^2}), \quad \ell \geq 1,$$

with the convention $g_\beta(0) = 1$. Using the i.i.d. assumption on the charges, the annealed partition function equals

$$(2.96) \quad \mathbb{E} \bar{Z}_n^{\beta, \omega} = \mathbb{E} \left[\prod_{x \in \mathbb{N}} g_\beta(\ell_n(x)) \right], \quad \text{with } \ell_n(x) := \text{card}\{1 \leq k \leq n : S_k = x\}.$$

Using the geometric distribution of the time spent on each visited vertex, we readily obtain

$$(2.97) \quad \mathbb{E} \bar{Z}_n^{\beta, \omega} = 2 \sum_{k=1}^n \sum_{\substack{\ell_1, \dots, \ell_k \geq 1 \\ \ell_1 + \dots + \ell_k = n}} \prod_{i=1}^k g_\beta(\ell_i) 2^{-\ell_i},$$

from which we deduce the annealed grand canonical partition function

$$(2.98) \quad \sum_{n \geq 1} \mathbb{E} \bar{Z}_n^{\beta, \omega} e^{fn} = 2 \sum_{k \geq 1} \left[\sum_{\ell \geq 1} (e^f/2)^\ell g_\beta(\ell) \right]^k.$$

The extra factor 2 in front of the sums in (2.97) and (2.98) comes from the fact that $S_1 = 1$. Thus, the annealed free energy, which could alternatively be defined as

$$(2.99) \quad \bar{F}_{\text{ann}}(\beta) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} \bar{Z}_n^{\beta, \omega},$$

has the variational representation displayed in (2.94). The first inequality therein follows from Theorem 2.10 and the standard Jensen bound:

$$(2.100) \quad \mathbb{E} \log \bar{Z}_n^{\beta, \omega} \leq \log \mathbb{E} \bar{Z}_n^{\beta, \omega}.$$

The fact that the annealed free energy is non-positive clearly follows from (2.99). The fact that it is bounded from below by $(-\log 2)$ follows by checking that

$$(2.101) \quad \sum_{\ell \geq 1} (e^{\log 2}/2)^\ell g_\beta(\ell) = \sum_{\ell \geq 1} g_\beta(\ell) \geq \sum_{\ell \in 2\mathbb{N}} \mathbb{P}(\Omega_\ell = 0) = +\infty,$$

since $\mathbb{P}(\Omega_\ell = 0) \sim (\text{cst.}) \ell^{-1/2}$ as $\ell \rightarrow \infty$ along even integers. \blacksquare

We first use Proposition 2.17 to derive a high-temperature upper bound.

Proposition 2.18 (Binary charges). *As $\beta \rightarrow 0$,*

$$(2.102) \quad \bar{F}_{\text{que}}(\beta) \leq \bar{F}_{\text{ann}}(\beta) \leq -\beta + 2\beta^2 + o(\beta^2).$$

Proposition 2.19 (standard Gaussian charges). *As $\beta \rightarrow 0$,*

$$(2.103) \quad \bar{F}_{\text{que}}(\beta) \leq \bar{F}_{\text{ann}}(\beta) \leq -\beta + 3\beta^2 + o(\beta^2).$$

Proof of Proposition 2.18. Recall the expression of the annealed free energy in (2.94). Replacing the $(e^f/2)$ therein by $z < 1$, using the inequality $e^{-y} \leq 1 - y + \frac{1}{2}y^2$ for every $y \geq 0$, and using Lemma A.1, we have

$$(2.104) \quad \sum_{\ell \geq 1} z^\ell \mathbb{E}(e^{-\beta\Omega_\ell^2}) \leq \sum_{\ell \geq 1} z^\ell [1 - (\beta + \beta^2)\ell + \frac{3}{2}\beta^2\ell^2].$$

Using Lemma A.2 with $\mathbf{c}_1(\beta) = -\beta - \beta^2$ and $\mathbf{c}_2(\beta) = \frac{3}{2}\beta^2$, we obtain

$$(2.105) \quad 2z^3 + (\mathbf{u}_1(\beta) - 5)z^2 + (\mathbf{u}_2(\beta) + 4)z - 1 < 0 \Rightarrow \sum_{\ell \geq 1} z^\ell \mathbb{E}(e^{-\beta\Omega_\ell^2}) < 1,$$

where

$$(2.106) \quad \mathbf{u}_1(\beta) = \beta + \frac{5}{2}\beta^2, \quad \mathbf{u}_2(\beta) = -\beta + \frac{1}{2}\beta^2.$$

Using Lemma A.3, we see that the condition on the l.h.s. of (3.7) is satisfied if

$$(2.107) \quad z < z(\beta) := z_0(\mathbf{u}_1(\beta), \mathbf{u}_2(\beta)) = \frac{1}{2} + \frac{1}{2}\beta - \frac{3}{4}\beta^2 + o(\beta^2).$$

Therefore,

$$(2.108) \quad \bar{F}_{\text{ann}}(\beta) \leq -\log(2z(\beta)) = -\beta + 2\beta^2 + o(\beta^2). \quad \blacksquare$$

Proof of Proposition 2.19. Same as Proposition 2.18 with $\mathbb{E}(\Omega_\ell^4) = 3\ell^2$, hence $\mathbf{c}_1(\beta) = -\beta$ and $\mathbf{c}_2(\beta) = \frac{3}{2}\beta^2$. \blacksquare

2.5. Low-temperature estimate on the free energy. We may also use Proposition 2.17 to derive *low-temperature* upper bounds.

Proposition 2.20 (Gaussian case). *As $\beta \rightarrow \infty$,*

$$(2.109) \quad \bar{F}_{\text{que}}(\beta) \leq \bar{F}_{\text{ann}}(\beta) = -\log 2 + \frac{\pi}{2\beta}(1 + o(1)),$$

Proof of Proposition 2.20. The inequality has already been proved so we focus on the low-temperature expansion. Recall the expression of the annealed free energy in (2.94). In the case of standard Gaussian charges, we can explicitly compute

$$(2.110) \quad \mathbb{E}(e^{-\beta\Omega_\ell^2}) = \frac{1}{\sqrt{1 + 2\beta\ell}}.$$

Considering $f(\beta) := \log 2 - C/\beta$, with $C > 0$, we obtain

$$(2.111) \quad \sum_{\ell \geq 1} (e^{f(\beta)}/2)^\ell \mathbb{E}(e^{-\beta\Omega_\ell^2}) = \sum_{\ell \geq 1} \frac{e^{-C\ell/\beta}}{\sqrt{1 + 2\beta\ell}}.$$

Dropping one in the square root above yields the following upper bound:

$$(2.112) \quad \sum_{\ell \geq 1} (e^{f(\beta)}/2)^\ell \mathbb{E}(e^{-\beta\Omega_\ell^2}) \leq \sum_{\ell \geq 1} \frac{e^{-C\ell/\beta}}{\sqrt{2\beta\ell}} = \frac{1}{\beta} \sum_{\ell \geq 1} \frac{e^{-C\ell/\beta}}{\sqrt{2\ell/\beta}},$$

which converges in the large β limit, by a Riemann approximation, to

$$(2.113) \quad \int_0^\infty (2x)^{-1/2} e^{-Cx} dx = \frac{1}{\sqrt{C}} \int_0^\infty (2x)^{-1/2} e^{-x} dx \stackrel{(y^2=2x)}{=} \sqrt{\frac{\pi}{2C}}.$$

The upper bound that we have just used can be complemented by the following lower bound:

$$(2.114) \quad \frac{e^{-C\ell/\beta}}{\sqrt{1+2\beta\ell}} = \frac{e^{-C\ell/\beta}}{\sqrt{2\beta\ell}} \times \frac{1}{\sqrt{1+(2\beta\ell)^{-1}}} \geq \frac{e^{-C\ell/\beta}}{\sqrt{2\beta\ell}} \times \frac{1}{\sqrt{1+(2\beta)^{-1}}}, \quad \ell \geq 1,$$

which allows to complete the proof. \blacksquare

Remark 2.21 (Binary case). *The annealed upper bound obtained in the case of binary charges is quite rough. Indeed, one can show that*

$$(2.115) \quad \lim_{\beta \rightarrow \infty} \bar{F}_{\text{ann}}(\beta) = -\frac{\log 3}{2} > -\log 2.$$

Existence follows from monotonicity. In order to obtain the value of the limit, we first observe that

$$(2.116) \quad 0 \leq \sum_{\ell \geq 1} z^\ell \mathbb{E}(e^{-\beta\Omega_\ell^2}) - \sum_{\ell \geq 1} z^{2\ell} \mathbb{P}(\Omega_{2\ell} = 0) \leq \frac{e^{-\beta} z}{1-z}, \quad 0 \leq z < 1.$$

Indeed, by splitting the leftmost sum according to the parity of ℓ , we obtain:

$$(2.117) \quad \begin{aligned} 0 &\leq \sum_{\substack{\ell \geq 1 \\ \ell \in 2\mathbb{Z}}} z^\ell \mathbb{E}(e^{-\beta\Omega_\ell^2}) - \sum_{\substack{\ell \geq 1 \\ \ell \in 2\mathbb{Z}}} z^\ell \mathbb{P}(\Omega_\ell = 0) \leq e^{-\beta} \sum_{\substack{\ell \geq 1 \\ \ell \in 2\mathbb{Z}}} z^\ell, \\ 0 &\leq \sum_{\substack{\ell \geq 1 \\ \ell \in 2\mathbb{Z}+1}} z^\ell \mathbb{E}(e^{-\beta\Omega_\ell^2}) \leq e^{-\beta} \sum_{\substack{\ell \geq 1 \\ \ell \in 2\mathbb{Z}+1}} z^\ell, \end{aligned}$$

from which (2.116) readily follows. The rightmost sum in (2.116) can be computed explicitly, see e.g. [21, Chapter I, E2]

$$(2.118) \quad \sum_{\ell \geq 1} z^{2\ell} \mathbb{P}(\Omega_{2\ell} = 0) = \frac{1}{\sqrt{1-z^2}} - 1, \quad |z| < 1.$$

Combining the above expression with (2.116) and (2.94) gives (2.115). The main contribution to the annealed partition function, in the large β limit, is thus given by charge distributions such that $\Omega_n(x) = 0$ for every $x \in \{S_1, \dots, S_n\}$. This indicates that any attempt to improve the annealed bound by means of constrained annealing with a linear potential, that is

$$(2.119) \quad \mathbb{E} \log \bar{Z}_n^{\beta, \omega} \leq \log \mathbb{E}[\bar{Z}_n^{\beta, \omega} \exp(\lambda \Omega_n)], \quad \lambda \in \mathbb{R},$$

(also known as first order Morita approximation) is bound to fail. See e.g. [9] for a reference on constrained annealing in the context of pinning

3. CORRECTIONS TO THE UNDIRECTED MODEL

In this section we correct two propositions from [8, Appendix D] concerning the *undirected* quenched charged polymer. The statement of the first proposition therein (Proposition D.1) is left unchanged, although we stress that the result holds in any dimension. The proof however must be corrected. In the second proposition (Proposition D.2), the sufficient condition for ballistic behaviour in dimension one is amended.

Although this section is independent of the rest of the paper, we prefer to stick to our notation (that is anyway not far from the one used in [8]) in order to make the paper more self-contained. The two major differences with the directed model from Section 2 are the following:

- \mathbb{P} denotes the law of *simple random walk* $S = (S_i)_{i \in \mathbb{N}_0}$ (started at the origin) on \mathbb{Z}^d , i.e. $(S_i - S_{i-1})_{i \in \mathbb{N}}$ is a sequence of i.i.d. random variables uniformly distributed on the $2d$ unit vectors.
- \mathbb{P} denotes the law of i.i.d. real-valued random variables $\omega = (\omega_i)_{i \in \mathbb{N}}$ with zero mean, unit variance and finite exponential moments, i.e. $M(\delta) := \mathbb{E}(e^{\delta \omega_1}) < \infty$ for every $\delta \geq 0$. Here, we also allow *biased* charge distributions: for every $\delta \geq 0$, we let \mathbb{P}_δ be the *tilted* probability measure uniquely defined by the property:

$$(3.1) \quad \mathbb{P}_\delta(A) = \mathbb{E}(e^{\delta \Omega_n - n \log M(\delta)} \mathbf{1}_A), \quad \forall n \in \mathbb{N}, \quad \forall A \in \sigma(\omega_1, \dots, \omega_n).$$

Note that the random variables $(\omega_i)_{i \in \mathbb{N}}$ remain i.i.d. under \mathbb{P}_δ and that $\mathbb{P}_0 = \mathbb{P}$.

In addition, we let

$$(3.2) \quad R_n(S) := \text{card}\{S_1, \dots, S_n\}$$

be the number of distinct vertices visited by the random walk between time 1 and time n .

The proposition below corresponds to [8, Proposition D.1]. The lower bound in the proof must be corrected. Also, we stress that the result is valid in any dimension.

Proposition 3.1. *Let $d \geq 1$. Suppose that $\delta, \beta \in (0, \infty)$. Then there exist $c_1, c_2 > 0$ (depending on δ, β) such that, for \mathbb{P}_δ -a.e. ω ,*

$$(3.3) \quad \bar{\mathbb{P}}_n^{\beta, \omega}(R_n(S) \leq c_1 n) \leq e^{-c_2 n + o(n)}.$$

Proof of Proposition 3.1. Let π be the one-sided path that takes right-steps only, i.e., $\pi_i = (i, 0, \dots, 0)$ for $i \in \mathbb{N}_0$. Let us denote by $\bar{H}_n^\omega(S)$ the Hamiltonian in (2.7) and estimate

$$(3.4) \quad \bar{Z}_n^{\beta, \omega} \geq \mathbb{E} \left[e^{-\beta \bar{H}_n^\omega(S)} \mathbf{1}_{\{S_i = \pi_i \forall 1 \leq i \leq n\}} \right] = \left(\frac{1}{2d}\right)^n e^{-\beta \sum_{i=1}^n \omega_i^2} = \left(\frac{1}{2d}\right)^n e^{-\beta \mathbb{E}_\delta(\omega_1^2)n + o(n)},$$

where the $o(\cdot)$ holds \mathbb{P}_δ -a.s. by the law of large numbers. Moreover, by Jensen's inequality we have (recall (2.8), (2.9) and the definition of ℓ_n in (2.43))

$$(3.5) \quad \begin{aligned} \bar{H}_n^\omega(S) &= \sum_{\substack{x \in \mathbb{Z}: \\ \ell_n(x) > 0}} \left(\sum_{i=1}^n \omega_i \mathbf{1}_{\{S_i = x\}} \right)^2 = R_n(S) \left[\frac{1}{R_n(S)} \sum_{\substack{x \in \mathbb{Z}: \\ \ell_n(x) > 0}} \left(\sum_{i=1}^n \omega_i \mathbf{1}_{\{S_i = x\}} \right)^2 \right] \\ &\geq R_n(S) \left(\frac{1}{R_n(S)} \sum_{\substack{x \in \mathbb{Z}: \\ \ell_n(x) > 0}} \sum_{i=1}^n \omega_i \mathbf{1}_{\{S_i = x\}} \right)^2 = \frac{1}{R_n(S)} \Omega_n^2. \end{aligned}$$

Combining (3.4–3.5), we obtain

$$(3.6) \quad \begin{aligned} \bar{\mathbb{P}}_n^{\beta, \omega}(R_n(S) \leq c_1 n) &\leq e^{\beta[\mathbb{E}_\delta(\omega_1^2) + o(1)]n} (2d)^n \mathbb{E} \left[\exp \left\{ -\frac{\beta}{R_n(S)} \Omega_n^2 \right\} \mathbf{1}_{\{R_n(S) \leq c_1 n\}} \right] \\ &\leq \exp \left\{ -\beta n \left[\frac{1}{c_1 n^2} \Omega_n^2 - \mathbb{E}_\delta(\omega_1^2) - \frac{\log(2d)}{\beta} + o(1) \right] \right\}. \end{aligned}$$

Note that $\mathbb{E}_\delta(\omega_1^2) = m(\delta)^2 + m'(\delta)$, where $m(\delta) := (\log M)'(\delta)$. By the strong law of large numbers for ω , we have $\lim_{n \rightarrow \infty} n^{-1} \Omega_n = m(\delta) > 0$ for \mathbb{P}_δ -a.e. ω , and so the term between

square brackets equals $c_3[1 + o(1)]$ with $c_3 = (\frac{1}{c_1} - 1)m(\delta)^2 - m'(\delta) - \frac{\log(2d)}{\beta}$. Therefore, by choosing $c_1 > 0$ small enough so that $c_3 > 0$, we get (3.3) with $c_2 = \beta c_3$. ■

The following proposition corrects [8, Proposition D.2]. The correction comes from the changes in the proof of Proposition 3.1.

Proposition 3.2. *Assume that $d = 1$ and that $\beta, \delta \in (0, \infty)$ satisfy*

$$(3.7) \quad m(\delta)^2 - m'(\delta) > \frac{\log 2}{\beta}.$$

Then, there exists $\varepsilon = \varepsilon(\delta, \beta) > 0$ such that:

$$(3.8) \quad \lim_{n \rightarrow \infty} \bar{\mathbb{P}}_n^{\beta, \omega}(n^{-1}S_n > \varepsilon \mid S_n > 0) = 1.$$

Remark 3.3. *When $\mathbb{P}(\omega_1 \in \cdot) = \mathcal{N}(0, 1)$ then $m(\delta) = \delta$ and $m'(\delta) = 1$, so that the condition in (3.7) is equivalent to $\delta^2 > 1 + (\log 2)/\beta$ (that is a non-empty condition). When $\mathbb{P}(\omega_1 \in \cdot)$ is the uniform probability measure on $\{-1, 1\}$ then $m(\delta) = \tanh(\delta)$ and $m'(\delta) = 1 - \tanh(\delta)^2$, so that the condition in (3.7) is equivalent to $2 \tanh(\delta)^2 > 1 + (\log 2)/\beta$. The latter condition is non-empty if and only if $\beta > \log 2$. A similar remark can be made for all bounded charge distributions.*

Proof of Proposition 3.2. Recall the value of c_3 in the proof of Proposition 3.1. If (3.7) is satisfied then we can choose $c_1 > \frac{1}{2}$ in Proposition 3.1 and use the inequality

$$(3.9) \quad \frac{|\{x \in \mathbb{Z}: \ell_n(x) = 1\}|}{n} \geq \frac{2R_n(S)}{n} - 1$$

to conclude that a positive fraction of the sites are visited precisely once. Consequently, if the polymer chain chooses to go to the right, then S_n/n has a strictly positive lim inf. ■

4. CONCLUSION AND PERSPECTIVES

For the *directed* charged polymer with quenched *centered* charges, we have reviewed and provided detailed proofs of past results concerning the freezing transition predicted by Derrida, Griffiths and Higgs. We showed that the quenched free energy enjoys a variational representation based on a quenched (or conditional) large deviation principle, much like pinning and copolymer models. Lower bounds are derived from this variational formula, while upper bounds are obtained through annealing. It is however not clear to us whether the freezing transition can be read from the free energy. Nevertheless, we hope that the efforts to obtain sharper estimates will lead to a better understanding of the model.

For the *undirected* charged polymer with quenched *biased* charges, one can show that the number of visited vertices is proportional to a positive fraction of the chain, with large probability. This property may be upgraded to true ballistic behavior, at least for a large enough charge bias, with an extra condition on the temperature for some charge distributions.

Let us finally close the paper with a (non-exhaustive) list of unsettled issues.

- For the directed model:
 - (1) Settle Items (1) and (2) in Conjecture 2.1.
 - (2) Give a rigorous proof for the existence of a low temperature regime in the case of random centered charges, as in Item (3) in Conjecture 2.1.

- (3) What is the critical temperature for the di-block polymer from (2.18)? The best estimate so far is $\log(\frac{1+\sqrt{5}}{2}) \leq \beta_c^{\text{di-block}} \leq 2 \log 2$.
- (4) Is the inequality $\bar{F}_{\text{que}}(\beta) \geq -\log 2$ strict? Is it valid beyond the binary charge distribution? The strategy used in the proof of Item (2) in Proposition 2.11 could be adapted to other charge distributions (e.g. Gaussian charges) by folding the chain along the sequence of stopping times defined by $\sigma_0 = 0$ and

$$(4.1) \quad \sigma_i = \min\{k > \sigma_{i-1} : \Omega_k > \Omega_{\sigma_{i-1}}\}, \quad i \in \mathbb{N},$$

instead of (2.59) (assuming w.l.o.g. that $\Omega_n > 0$). However, the fact that we must stop at $i_n := \min\{i \in \mathbb{N} : \Omega_{\sigma_i} \geq \Omega_n\}$, that is not a stopping time, would probably lead to cumbersome technicalities.

- (5) Is there a minimizer for the variational problem in (2.67)? If so, is it unique?
- (6) Is the quenched free energy analytic as a function of the inverse temperature?
- (7) Find the value of the constant C in the high temperature expansion

$$(4.2) \quad \bar{F}_{\text{que}}(\beta) = -\beta + C\beta^2 + o(\beta^2), \quad \beta \rightarrow 0.$$

Is it universal? Our best estimate so far is $1/2 \leq C \leq 2$ for binary charges and $1/2 \leq C \leq 3$ for Gaussian charges.

- (8) Can the lower bound in Theorem 2.13 be improved by considering a more subtle strategy (looking beyond consecutive monomers)?
- (9) Can we prove that $\bar{F}_{\text{que}}(\beta) \sim -\log 2$ as $\beta \rightarrow \infty$ in the case of binary charges? An upper bound is missing.

- For the undirected model:

- (1) Can we prove ballistic behavior for *every* $\delta, \beta \in (0, \infty)$?
- (2) What can be said about the case of *neutral* charges ($\beta = 0$)?

APPENDIX A. TECHNICAL ESTIMATES

Lemma A.1. *In the case of $\{-1, 1\}$ -valued centered i.i.d. charges, for every $\ell \geq 1$,*

$$(A.1) \quad \mathbb{E}(\Omega_\ell^4) = 3\ell^2 - 2\ell.$$

Proof of Lemma A.1. We have

$$(A.2) \quad \mathbb{E}(\Omega_\ell^4) = \mathbb{E}\left(\left[\ell + 2 \sum_{1 \leq i < j \leq \ell} \omega_i \omega_j\right]^2\right) = \ell^2 + 4\mathbb{E}\left(\sum_{\substack{1 \leq i < j \leq \ell \\ 1 \leq u < v \leq \ell}} \omega_i \omega_j \omega_u \omega_v\right),$$

and the result follows by noticing that only the terms for which $i = u$ and $j = v$ give a nonzero (and actually unit) contribution to the last expectation. \blacksquare

Lemma A.2. *Let $|z| < 1$. Then,*

$$(A.3) \quad \sum_{\ell \geq 1} z^\ell (1 + c_1 \ell + c_2 \ell^2) < 1 \Leftrightarrow 2z^3 + (u_1 - 5)z^2 + (u_2 + 4)z - 1 < 0,$$

where

$$(A.4) \quad u_1 = c_2 - c_1, \quad u_2 = c_1 + c_2.$$

Proof of Lemma A.2. The sum in the l.h.s. of (A.3) equals

$$(A.5) \quad \sum_{\ell \geq 1} z^\ell + [c_1 + c_2] \sum_{\ell \geq 1} \ell z^\ell + c_2 \sum_{\ell \geq 1} \ell(\ell - 1)z^\ell = \frac{z}{1-z} + \frac{[c_1 + c_2]z}{(1-z)^2} + \frac{2c_2 z^2}{(1-z)^3}.$$

The lemma follows from a straightforward computation. \blacksquare

Lemma A.3. *For every $\mathbf{u} \in \mathbb{R}^2$ and $z \in \mathbb{R}$, let*

$$(A.6) \quad P_{\mathbf{u}}(z) = 2z^3 + (\mathbf{u}_1 - 5)z^2 + (\mathbf{u}_2 + 4)z - 1.$$

There exists a neighborhood of the origin $\mathcal{N} \subseteq \mathbb{R}^2$ and a function $\mathbf{u} \in \mathcal{N} \mapsto z_0(\mathbf{u})$ such that, for every $\mathbf{u} \in \mathcal{N}$ such that $\mathbf{u}_1 + \mathbf{u}_2 > 0$, $z_0(\mathbf{u})$ is the only root of $P_{\mathbf{u}}$ in $(-\infty, 1)$. Moreover,

$$(A.7) \quad z_0(\mathbf{u}) = \frac{1}{2} - \frac{\mathbf{u}_1}{2} - \mathbf{u}_2 + 2\mathbf{u}_1^2 + 6\mathbf{u}_2^2 + 7\mathbf{u}_1\mathbf{u}_2 + o(\|\mathbf{u}\|^2), \quad \mathbf{u} \rightarrow 0.$$

Proof of Lemma A.3. First, observe that

$$(A.8) \quad P_0(z) = (2z - 1)(z - 1)^2$$

admits $1/2$ as a simple root. The result follows from the Implicit Function Theorem and an explicit second-order Taylor expansion. Letting

$$(A.9) \quad z_0(\mathbf{u}) = \frac{1}{2} + \langle (Dz_0)(0), \mathbf{u} \rangle + \frac{1}{2} \langle \mathbf{u}, (D^2z_0)(0)\mathbf{u} \rangle + o(\|\mathbf{u}\|^2), \quad \mathbf{u} \rightarrow 0,$$

(with column vectors) we obtain

$$(A.10) \quad \begin{aligned} 0 &= P_{\mathbf{u}}(z_0(\mathbf{u})) = P_0(z_0(\mathbf{u})) + \mathbf{u}_1 z_0(\mathbf{u})^2 + \mathbf{u}_2 z_0(\mathbf{u}) \\ &= P_0'(\tfrac{1}{2})[z_0(\mathbf{u}) - \tfrac{1}{2}] + \tfrac{1}{2} P_0''(\tfrac{1}{2})[z_0(\mathbf{u}) - \tfrac{1}{2}]^2 \\ &\quad + \tfrac{1}{4} \mathbf{u}_1 + \mathbf{u}_1 [z_0(\mathbf{u}) - \tfrac{1}{2}] + \tfrac{1}{2} \mathbf{u}_2 + \mathbf{u}_2 [z_0(\mathbf{u}) - \tfrac{1}{2}] + o(\|\mathbf{u}\|^2) \\ &= \langle \tfrac{1}{2} (Dz_0)(0) + (\tfrac{1}{4}, \tfrac{1}{2})^\top, \mathbf{u} \rangle \\ &\quad + \langle \mathbf{u}, [\tfrac{1}{4} (D^2z_0)(0) - 2(Dz_0)(0)(Dz_0)(0)^\top + (1, 1)^\top (Dz_0)(0)^\top] \mathbf{u} \rangle + o(\|\mathbf{u}\|^2). \end{aligned}$$

Therefore,

$$(A.11) \quad (Dz_0)(0) = \begin{pmatrix} -1/2 \\ -1 \end{pmatrix}, \quad (D^2z_0)(0) = \begin{pmatrix} 4 & 7 \\ 7 & 12 \end{pmatrix}.$$

The restriction that $\mathbf{u}_1 + \mathbf{u}_2 > 0$ in the statement comes from the fact that $P_{\mathbf{u}}(1) = \mathbf{u}_1 + \mathbf{u}_2$ and allows us to verify that $z_0(\mathbf{u})$ is indeed the only root in $(-\infty, 1)$ rather than in a neighborhood of $1/2$. \blacksquare

APPENDIX B. HIGH-TEMPERATURE EXPANSION OF THE FINITE-VOLUME FREE ENERGY

In this section we provide the necessary details for the computation behind Remark 2.15.

Proposition B.1. *In addition to the ω_i 's being i.i.d. square integrable random variables with zero mean and unit variance, we assume that $\mathbb{E}(\omega_1^4) < \infty$. For every $n \in \mathbb{N}$, the finite-volume (averaged) quenched free energy has the following high-temperature ($\beta \rightarrow 0$) expansion:*

$$(B.1) \quad \frac{1}{n} \mathbb{E} \log \bar{Z}_n^{\beta, \omega} = -\beta + \mathbf{c}_n \beta^2 + o(\beta^2), \quad \text{with} \quad \lim_{n \rightarrow \infty} \mathbf{c}_n = 4/3.$$

Proof of Proposition B.1. Let $n \in \mathbb{N}$. Expanding the exponential in (2.8), we have

$$(B.2) \quad \bar{Z}_n^{\beta, \omega} = 1 - \beta \sum_{x \in \mathbb{N}} \mathbb{E}[\Omega_n(x)^2] + [1 + o(1)] \frac{\beta^2}{2} \sum_{x, y \in \mathbb{N}} \mathbb{E}[\Omega_n(x)^2 \Omega_n(y)^2].$$

We now apply the logarithm and expand it up to the second-order term. Using that

$$(B.3) \quad \sum_{x \in \mathbb{N}} \mathbb{E} \mathbb{E}[\Omega_n(x)^2] = \sum_{x \in \mathbb{N}} \ell_n(x) = n,$$

we find thereby:

$$(B.4) \quad \mathbb{E} \log \bar{Z}_n^{\beta, \omega} = -\beta n + [1 + o(1)] \frac{\beta^2}{2} C_n,$$

where

$$(B.5) \quad \begin{aligned} C_n &:= \sum_{x, y \in \mathbb{N}} \mathbb{E} \text{Cov}(\Omega_n(x)^2, \Omega_n(y)^2) \\ &= \sum_{x, y \in \mathbb{N}} \mathbb{E} \mathbb{E}^{\otimes 2}[\Omega_n(x)^2 \Omega_n(y)^2 - \Omega_n(x)^2 \tilde{\Omega}_n(y)^2], \end{aligned}$$

and \tilde{S} is a replica (i.e. independent copy) of S (but the fields Ω_n and $\tilde{\Omega}_n$ are built on a common sequence of charges ω). Recalling (2.9) and expanding all the squares in the above expression, C_n is found to be equal to:

$$(B.6) \quad \sum_{\substack{x, y \in \mathbb{N} \\ 1 \leq i, j, k, \ell \leq n}} \mathbb{E}(\omega_i \omega_j \omega_k \omega_\ell) [\mathbb{P}(S_i = S_j = x, S_k = S_\ell = y) - \mathbb{P}(S_i = S_j = x) \mathbb{P}(S_k = S_\ell = y)].$$

Due to our set of assumptions on the charge sequence ω , only the four following cases may contribute to the sum:

- (1) $i = j = k = \ell$;
- (2) $i = j$ and $k = \ell$, but $i \neq k$;
- (3) $i = k$ and $j = \ell$, but $i \neq j$;
- (4) $i = \ell$ and $j = k$, but $i \neq j$.

The reader may check that Cases (1) and (2) give a zero contribution (once we sum over x and y) while Cases (3) and (4) give the same contribution. Therefore, as $n \rightarrow \infty$,

$$(B.7) \quad C_n = 4 \sum_{1 \leq i < j \leq n} [\mathbb{P}(S_i = S_j) - \mathbb{P}(S_i = S_j)^2] \sim 4n \sum_{\ell \geq 1} [2^{-\ell} - 4^{-\ell}] = \frac{8n}{3},$$

which completes the proof. ■

REFERENCES

- [1] Q. Berger, F. Caravenna, J. Poisat, R. Sun, and N. Zygouras. The critical curves of the random pinning and copolymer models at weak coupling. *Comm. Math. Phys.*, 326(2):507–530, 2014.
- [2] Q. Berger, F. den Hollander, and J. Poisat. Annealed scaling for a charged polymer in dimensions two and higher. *J. Phys. A*, 51(5):054002, 37, 2018.
- [3] M. Birkner. Conditional large deviations for a sequence of words. *Stochastic Process. Appl.*, 118(5):703–729, 2008.
- [4] M. Birkner, A. Greven, and F. den Hollander. Quenched large deviation principle for words in a letter sequence. *Probab. Theory Related Fields*, 148(3-4):403–456, 2010.

- [5] M. Birkner, A. Greven, and F. d. Hollander. Correction to: Quenched large deviation principle for words in a letter sequence. *Probab. Theory Related Fields*, 187(1-2):523–569, 2023.
- [6] E. Bolthausen, F. den Hollander, and A. A. Opoku. A copolymer near a selective interface: variational characterization of the free energy. *Ann. Probab.*, 43(2):875–933, 2015.
- [7] F. Caravenna and F. den Hollander. Phase transitions for spatially extended pinning. *Probab. Theory Related Fields*, 181(1-3):329–375, 2021.
- [8] F. Caravenna, F. den Hollander, N. Pétrelis, and J. Poisat. Annealed scaling for a charged polymer. *Math. Phys. Anal. Geom.*, 19(1):Art. 2, 87, 2016.
- [9] F. Caravenna and G. Giacomin. On constrained annealed bounds for pinning and wetting models. *Electron. Comm. Probab.*, 10:179–189, 2005.
- [10] D. Cheliotis and F. den Hollander. Variational characterization of the critical curve for pinning of random polymers. *Ann. Probab.*, 41(3B):1767–1805, 2013.
- [11] F. Comets. *Directed polymers in random environments*, volume 2175 of *Lecture Notes in Mathematics*. Springer, Cham, 2017. Lecture notes from the 46th Probability Summer School held in Saint-Flour, 2016.
- [12] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
- [13] F. den Hollander. *Random polymers*, volume 1974 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. Lectures from the 37th Probability Summer School held in Saint-Flour, 2007.
- [14] F. den Hollander and A. A. Opoku. Copolymer with pinning: variational characterization of the phase diagram. *J. Stat. Phys.*, 152(5):846–893, 2013.
- [15] B. Derrida, R. Griffiths, and P. Higgs. A model of directed random walks with random selfinteractions. *Europhys. Lett.*, 18:361–366, 1992.
- [16] B. Derrida and P. G. Higgs. Low-temperature properties of directed walks with random self-interactions. *J. Phys. A*, 27(16):5485–5493, 1994.
- [17] G. Giacomin. *Random polymer models*. Imperial College Press, London, 2007.
- [18] G. Giacomin. *Disorder and critical phenomena through basic probability models*, volume 2025 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lecture notes from the 40th Probability Summer School held in Saint-Flour, 2010, École d’Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School].
- [19] Y. Kantor and M. Kardar. Polymers with random self-interactions. *Europhys. Lett.*, 14:421–426, 1991.
- [20] J.-C. Mourrat. On the delocalized phase of the random pinning model. In *Séminaire de Probabilités XLIV*, volume 2046 of *Lecture Notes in Math.*, pages 401–407. Springer, Heidelberg, 2012.
- [21] F. Spitzer. *Principles of random walk*. Springer-Verlag, New York-Heidelberg, second edition, 1976.