IMS Lecture Notes–Monograph Series
Optimality: The Third Erich L. Lehmann Symposium
Vol. 57 (2009) 237–251

DOI: 10.1214/09-LNMS5714

From Charged Polymers to Random Walk in Random Scenery

Xia Chen[1,*] and Davar Khoshnevisan[2,†]

University of Tennessee and University of Utah

Abstract: We prove that two seemingly-different models of random walk in random environment are generically quite close to one another. One model comes from statistical physics, and describes the behavior of a randomly-charged random polymer. The other model comes from probability theory, and was originally designed to describe a large family of asymptotically self-similar processes that have stationary increments.

Contents

1 Introduction and the Main Results 237
2 Preliminary Estimates 239
3 Proofs of the Main Results 245
Acknowledgements 250
References 251

1. Introduction and the Main Results

The principal goal of this article is to show that two apparently-disparate models—one from statistical physics of disorder media (Kantor and Kardar [9], Derrida et al. [5], Derrida and Higgs [6]) and one from probability theory (Kesten and Spitzer [10], Bolthausen [1])—are very close to one another.

In order to describe the model from statistical physics, let us suppose that $q := \{q_i\}_{i=1}^{\infty}$ is a collection of i.i.d. mean-zero random variables with finite variance $\sigma^2 > 0$. For technical reasons, we assume here and throughout that

$$\mu_6 := \mathrm{E}(q_1^6) < \infty. \tag{1.1}$$

In addition, we let $S := \{S_i\}_{i=0}^{\infty}$ denote a random walk on $\mathbf{Z}^d$ with $S_0 = 0$ that is independent from the collection q. We also rule out the trivial case that S_1 has only one possible value.

The object of interest to us is the random quantity

$$H_n := \sum\sum_{1\le i<j\le n} q_i q_j \mathbf{1}_{\{S_i=S_j\}}. \tag{1.2}$$

[1]Department of Mathematics, University of Tennessee, Knoxville, TN 37996-1300, e-mail: xchen@math.utk.edu
*Research supported in part by NSF grant DMS-0704024.
[2]Department of Mathematics, University of Utah, Salt Lake City, UT 84112-0090, e-mail: davar@math.utah.edu
†Research supported in part by NSF grant DMS-0706728.
AMS 2000 subject classifications: Primary 60K35; secondary 60K37.
Keywords and phrases: polymer measures, random walk in random scenery.

In statistical physics, H_n denotes a random Hamiltonian of spin-glass type that is used to build Gibbsian polymer measures. The q_i's are random charges, and each realization of S corresponds to a possible polymer path; see the paper by Kantor and Kardar [9], its subsequent variations by Derrida et al. [5, 6] and Wittmer et al. [17], and its predecessos by Garel and Orland [7] and Obukhov [14]. The resulting Gibbs measure then corresponds to a model for "random walk in random environment." Although we do not consider continuous processes here, the continuum-limit analogue of H_n has also been studied in the literature (Buffet and Pulé [2], Martìnez and Petritis [13]).

Kesten and Spitzer [10] introduced a different model for "random walk in random environment," which they call *random walk in random scenery.*[1] We can describe that model as follows: Let $Z := \{Z(x)\}_{x\in\mathbf{Z}^d}$ denote a collection of i.i.d. random variables, with the same common distribution as q_1, and independent of S. Define

$$W_n := \sum_{i=1}^{n} Z(S_i). \tag{1.3}$$

The process $W := \{W_n\}_{n=0}^{\infty}$ is called *random walk in random scenery*, and can be thought of as follows: We fix a realization of the d-dimensional random field Z—the "scenery"—and then run an independent walk S on $\mathbf{Z}^d$. At time j, the walk is at S_j; we sample the scenery at that point. This yields $Z(S_j)$, which is then used as the increment of the process W at time j.

Our goal is to make precise the assertion that if n is large, then

$$H_n \approx \gamma^{1/2} \cdot W_n \quad \text{in distribution,} \tag{1.4}$$

where

$$\gamma := \begin{cases} 1 & \text{if } S \text{ is recurrent,} \\ \sum_{k=1}^{\infty} \mathrm{P}\{S_k = 0\} & \text{if } S \text{ is transient.} \end{cases} \tag{1.5}$$

Our derivation is based on a classification of recurrence vs. transience for random walks that appears to be new. This classification [Theorem 2.4] might be of independent interest.

We can better understand (1.4) by considering separately the cases that S is transient versus recurrent. The former case is simpler to describe, and appears next.

Theorem 1.1. *If S is transient, then*

$$\frac{W_n}{n^{1/2}} \xrightarrow{\mathcal{D}} N(0\,,\sigma^2) \quad \textit{and} \quad \frac{H_n}{n^{1/2}} \xrightarrow{\mathcal{D}} N(0\,,\gamma\sigma^2). \tag{1.6}$$

Kesten and Spitzer [10] proved the assertion about W_n under more restrictive conditions on S. Similarly, Chen [3] proved the statement about H_n under more hypotheses.

Before we can describe the remaining [and more interesting] recurrent case, we define

$$a_n := \left(n \sum_{k=0}^{n} \mathrm{P}\{S_k = 0\} \right)^{1/2}. \tag{1.7}$$

It is well known (Polya [15], Chung and Fuchs [4]) that S is recurrent if and only if $a_n/n^{1/2} \to \infty$ as $n \to \infty$.

[1] Kesten and Spitzer ascribe the terminology to Paul Shields.

Theorem 1.2. *If S is recurrent, then for all bounded continuous functions $f : \mathbf{R}^d \to \mathbf{R}$,*

$$\mathrm{E}\left[f\left(\frac{W_n}{a_n}\right)\right] = \mathrm{E}\left[f\left(\frac{H_n}{a_n}\right)\right] + o(1), \tag{1.8}$$

where $o(1)$ converges to zero as $n \to \infty$. Moreover, both $\{W_n/a_n\}_{n\geq 1}$ and $\{H_n/a_n\}_{n\geq 1}$ are tight.

We demonstrate Theorems 1.1 and 1.2 by using a variant of the replacement method of Liapounov [11] [pp. 362–364]; this method was rediscovered later by Lindeberg [12], who used it to prove his famous central limit theorem for triangular arrays of random variables.

It can be proved that when S is in the domain of attraction of a stable law, W_n/a_n converges in distribution to an explicit law (Kesten and Spitzer [10], Bolthausen [1]). Consequently, H_n/a_n converges in distribution to the same law in that case. This fact was proved earlier by Chen [3] under further [mild] conditions on S and q_1.

We conclude the introduction by describing the growth of a_n under natural conditions on S.

Remark 1.3. Suppose S is strongly aperiodic, mean zero, and finite second moments, with a nonsingular covariance matrix. Then, S is transient iff $d \geq 3$, and by the local central limit theorem, as $n \to \infty$,

$$\sum_{k=1}^{n} \mathrm{P}\{S_k = 0\} \sim \text{const} \times \begin{cases} n^{1/2} & \text{if } d = 1, \\ \log n & \text{if } d = 2. \end{cases} \tag{1.9}$$

See, for example (Spitzer [16] [**P9** on p. 75]). Consequently,

$$a_n \sim \text{const} \times \begin{cases} n^{3/4} & \text{if } d = 1, \\ (n \log n)^{1/2} & \text{if } d = 2. \end{cases} \tag{1.10}$$

This agrees with the normalization of Kesten and Spitzer [10] when $d = 1$, and Bolthausen [1] when $d = 2$.

2. Preliminary Estimates

Consider the local times of S defined by

$$L_n^x := \sum_{i=1}^{n} \mathbf{1}_{\{S_i = x\}}. \tag{2.1}$$

A little thought shows that the random walk in random scenery can be represented compactly as

$$W_n = \sum_{x \in \mathbf{Z}^d} Z(x) L_n^x. \tag{2.2}$$

There is also a nice way to write the random Hamiltonian H_n in local-time terms. Consider the "level sets,"

$$\mathcal{L}_n^x := \{i \in \{1, \ldots, n\} : \; S_i = x\}. \tag{2.3}$$

It is manifest that if $j \in \{2, \ldots, n\}$, then $L_j^x > L_{j-1}^x$ if and only if $j \in \mathcal{L}_n^x$. Thus, we can write

$$
\begin{aligned}
H_n &= \frac{1}{2}\left(\sum_{x\in\mathbf{Z}^d}\left|\sum_{i=1}^n q_i \mathbf{1}_{\{S_i=x\}}\right|^2 - \sum_{i=1}^n q_i^2\right) \\
&= \sum_{x\in\mathbf{Z}^d} h_n^x,
\end{aligned} \tag{2.4}
$$

where

$$
h_n^x := \frac{1}{2}\left(\left|\sum_{i\in\mathcal{L}_n^x} q_i\right|^2 - \sum_{i\in\mathcal{L}_n^x} q_i^2\right). \tag{2.5}
$$

We denote by $\widehat{\mathrm{P}}$ the conditional measure, given the entire process S; $\widehat{\mathrm{E}}$ denotes the corresponding expectation operator. The following is borrowed from Chen [3] [Lemma 2.1].

Lemma 2.1. *Choose and fix some integer $n \geq 1$. Then, $\{h_n^x\}_{x\in\mathbf{Z}^d}$ is a collection of independent random variables under $\widehat{\mathrm{P}}$, and*

$$
\widehat{\mathrm{E}} h_n^x = 0 \quad \textit{and} \quad \widehat{\mathrm{E}}\left(|h_n^x|^2\right) = \frac{\sigma^2}{2} L_n^x\left(L_n^x - 1\right) \qquad \mathrm{P}\text{-}a.s. \tag{2.6}
$$

Moreover, there exists a nonrandom positive and finite constant $C = C(\sigma)$ such that for all $n \geq 1$ and $x \in \mathbf{Z}^d$,

$$
\widehat{\mathrm{E}}\left(|h_n^x|^3\right) \leq C\mu_6\left|L_n^x\left(L_n^x - 1\right)\right|^{3/2} \qquad \mathrm{P}\text{-}a.s. \tag{2.7}
$$

Next we develop some local-time computations.

Lemma 2.2. *For all $n \geq 1$,*

$$
\sum_{x\in\mathbf{Z}^d} \mathrm{E} L_n^x = n \quad \textit{and} \quad \sum_{x\in\mathbf{Z}^d} \mathrm{E}\left(|L_n^x|^2\right) = n + 2\sum_{k=1}^{n-1}(n-k)\mathrm{P}\{S_k = 0\}. \tag{2.8}
$$

Moreover, for all integers $k \geq 1$,

$$
\sum_{x\in\mathbf{Z}^d} \mathrm{E}\left(|L_n^x|^k\right) \leq k!\, n\left|\sum_{j=0}^n \mathrm{P}\{S_j = 0\}\right|^{k-1}. \tag{2.9}
$$

Proof. Since $\mathrm{E}L_n^x = \sum_{j=1}^n \mathrm{P}\{S_j = x\}$ and $\sum_{x\in\mathbf{Z}^d}\mathrm{P}\{S_j = x\} = 1$, we have $\sum_x \mathrm{E}L_n^x = n$. For the second-moment formula we write

$$
\begin{aligned}
\mathrm{E}\left(|L_n^x|^2\right) &= \sum_{1\leq i\leq n} \mathrm{P}\{S_i = x\} + 2\sum\sum_{1\leq i<j\leq n} \mathrm{P}\{S_i = S_j = x\} \\
&= \sum_{1\leq i\leq n} \mathrm{P}\{S_i = x\} + 2\sum\sum_{1\leq i<j\leq n} \mathrm{P}\{S_i = x\}\mathrm{P}\{S_{j-i} = 0\}.
\end{aligned} \tag{2.10}
$$

We can sum this expression over all $x \in \mathbf{Z}^d$ to find that

$$
\sum_{x\in\mathbf{Z}^d} \mathrm{E}\left(|L_n^x|^2\right) = n + 2\sum\sum_{1\leq i<j\leq n} \mathrm{P}\{S_{j-i} = 0\}. \tag{2.11}
$$

This readily implies the second-moment formula. Similarly, we write

$$\begin{aligned}
&\mathrm{E}\left(|L_n^x|^k\right)\\
&\le k! \sum\cdots\sum_{1\le i_1\le\cdots\le i_k\le n} \mathrm{P}\{S_{i_1}=\cdots=S_{i_k}=x\}\\
&= k! \sum\cdots\sum_{1\le i_1\le\cdots\le i_k\le n} \mathrm{P}\{S_{i_1}=x\}\mathrm{P}\{S_{i_2-i_1}=0\}\cdots\mathrm{P}\{S_{i_k-i_{k-1}}=0\}\\
&\le k! \sum_{i=1}^n \mathrm{P}\{S_i=x\}\cdot\left|\sum_{j=1}^n \mathrm{P}\{S_j=0\}\right|^{k-1}.
\end{aligned}\tag{2.12}$$

Add over all $x\in\mathbf{Z}^d$ to finish. □

Our next lemma provides the first step in a classification of recurrence [versus transience] for random walks.

Lemma 2.3. *It is always the case that*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{x\in\mathbf{Z}^d}\mathrm{E}\left(|L_n^x|^2\right)=1+2\sum_{k=1}^\infty \mathrm{P}\{S_k=0\}.\tag{2.13}$$

Proof. Thanks to Lemma 2.2, for all $n\ge 1$,

$$\frac{1}{n}\sum_{x\in\mathbf{Z}^d}\mathrm{E}\left(|L_n^x|^2\right)=1+2\sum_{k=1}^{n-1}\left(1-\frac{k}{n}\right)\mathrm{P}\{S_k=0\}.\tag{2.14}$$

If S is transient, then the monotone convergence theorem ensures that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{x\in\mathbf{Z}^d}\mathrm{E}\left(|L_n^x|^2\right)=1+2\sum_{k=1}^\infty \mathrm{P}\{S_k=0\}.\tag{2.15}$$

This proves the lemma in the transient case.

When S is recurrent, we note that (2.14) readily implies that for all integers $m\ge 2$,

$$\begin{aligned}
\liminf_{n\to\infty}\frac{1}{n}\sum_{x\in\mathbf{Z}^d}\mathrm{E}\left(|L_n^x|^2\right)&\ge 1+2\sum_{k=1}^{m-1}\left(1-\frac{k}{m}\right)\mathrm{P}\{S_k=0\}\\
&\ge 1+\sum_{1\le k\le m/2}\mathrm{P}\{S_k=0\}.
\end{aligned}\tag{2.16}$$

Let $m\uparrow\infty$ to deduce the lemma. □

Next we "remove the expectation" from the statement of Lemma 2.3.

Theorem 2.4. *As $n\to\infty$,*

$$\frac{1}{n}\sum_{x\in\mathbf{Z}^d}(L_n^x)^2\to 1+2\sum_{k=1}^\infty \mathrm{P}\{S_k=0\}\quad \textit{in probability.}\tag{2.17}$$

Remark 2.5. The quantity $I_n := \sum_{x\in\mathbf{Z}^d}(L_n^x)^2$ is the socalled *self-intersection local time* of the walk S. This terminology stems from the following elementary calculation: For all integers $n \geq 1$,

$$I_n = \sum\sum_{1\leq i,j\leq n} \mathbf{1}_{\{S_j=S_i\}}. \tag{2.18}$$

Consequently, Theorem 2.4 implies that a random walk S on $\mathbf{Z}^d$ is recurrent if and only if its self-intersection local time satisfies $I_n/n \to \infty$ in probability.

Remark 2.6. Nadine Guillotin–Plantard has kindly pointed out to us that the mode of convergence in Theorem 2.4 can be strengthened to almost-sure convergence. This requires a direct subadditivity argument (Guillotin–Plantard [8]). It follows also from the estimates that follow, together with a classical blocking argument, which we skip.

Proof. First we study the case that $\{S_i\}_{i=0}^\infty$ is transient.

Define

$$Q_n := \sum\sum_{1\leq i<j\leq n} \mathbf{1}_{\{S_i=S_j\}}. \tag{2.19}$$

Then it is not too difficult to see that

$$\sum_{x\in\mathbf{Z}^d} (L_n^x)^2 = 2Q_n + n \qquad \text{for all } n \geq 1. \tag{2.20}$$

This follows immediately from (2.18), for example. Therefore, it suffices to prove that, under the assumption of transience,

$$\frac{Q_k}{k} \to \sum_{j=1}^\infty \mathrm{P}\{S_j = 0\} \quad \text{in probability as } k \to \infty. \tag{2.21}$$

Lemma 2.3 and (2.20) together imply that

$$\lim_{k\to\infty} \frac{\mathrm{E}Q_k}{k} = \sum_{j=1}^\infty \mathrm{P}\{S_j = 0\}. \tag{2.22}$$

Hence, it suffices to prove that $\operatorname{Var} Q_n = o(n^2)$ as $n \to \infty$. In some cases, this can be done by making an explicit [though hard] estimate for $\operatorname{Var} Q_n$; see, for instance, (Chen [3] [Lemma 5.1]), and also the technique employed in the proof of Lemma 2.4 of Bolthausen [1]. Here, we opt for a more general approach that is simpler, though it is a little more circuitous. Namely, in rough terms, we write Q_n as $Q_n^{(1)} + Q_n^{(2)}$, where $\mathrm{E}Q_n^{(1)} = o(n)$, and $\operatorname{Var} Q_n^{(2)} = o(n^2)$. Moreover, we will soon see that $Q_n^{(1)}, Q_n^{(2)} \geq 0$, and this suffices to complete the proof.

For all $m := m_n \in \{1, \ldots, n-1\}$ we write

$$Q_n = Q_n^{1,m} + Q_n^{2,m}, \tag{2.23}$$

where

$$Q_n^{1,m} := \sum\sum_{\substack{1\leq i<j\leq n:\\ j\geq i+m}} \mathbf{1}_{\{S_i=S_j\}} \quad \text{and} \quad Q_n^{2,m} := \sum\sum_{\substack{1\leq i<j\leq n:\\ j<i+m}} \mathbf{1}_{\{S_i=S_j\}}. \tag{2.24}$$

Because $n > m$, we have

$$\mathrm{E}Q_n^{1,m} \leq n \sum_{k=m}^{\infty} \mathrm{P}\{S_k = 0\}. \tag{2.25}$$

We estimate the variance of $Q_n^{2,m}$ next. We do this by first making an observation.

Throughout the remainder of this proof, define for all subsets Γ of $\mathbf{N}^2$,

$$\Upsilon(\Gamma) := \sum\sum_{(i,j)\in\Gamma} \mathbf{1}_{\{S_i = S_j\}}. \tag{2.26}$$

Supppose $\Gamma_1, \Gamma_2, \ldots, \Gamma_\nu$ are finite disjoint sets in $\mathbf{N}^2$, with common cardinality, and the added property that whenever $1 \leq a < b \leq \nu$, we have $\Gamma_a < \Gamma_b$ in the sense that $i < k$ and $j < l$ for all $(i\,, j) \in \Gamma_a$ and $(k\,, l) \in \Gamma_b$. Then, it follows that

$$\{\Upsilon(\Gamma_\nu)\}_{\mu=1}^{\nu} \quad \text{is an i.i.d. sequence.} \tag{2.27}$$

For all integers $p \geq 0$ define

$$\begin{aligned} B_p^m &:= \left\{(i\,, j) \in \mathbf{N}^2 :\, (p-1)m < i < j \leq pm\right\}, \\ W_p^m &:= \left\{(i\,, j) \in \mathbf{N}^2 :\, (p-1)m < i \leq pm < j \leq (p+1)m\right\}. \end{aligned} \tag{2.28}$$

In Figure 1, $\{B_p^m\}_{p=1}^{\infty}$ denotes the collection black and $\{W_p^m\}_{p=1}^{\infty}$ the white triangles that are inside the slanted strip.

We may write

$$Q_{(n-1)m}^{2,m} = \sum_{p=1}^{n-1} \Upsilon(B_p^m) + \sum_{p=1}^{n-1} \Upsilon(W_p^m). \tag{2.29}$$

Consequently,

$$\operatorname{Var} Q_{(n-1)m}^{2,m} \leq 2\operatorname{Var} \sum_{p=1}^{n-1} \Upsilon(B_p^m) + 2\operatorname{Var} \sum_{p=1}^{n-1} \Upsilon(W_p^m). \tag{2.30}$$

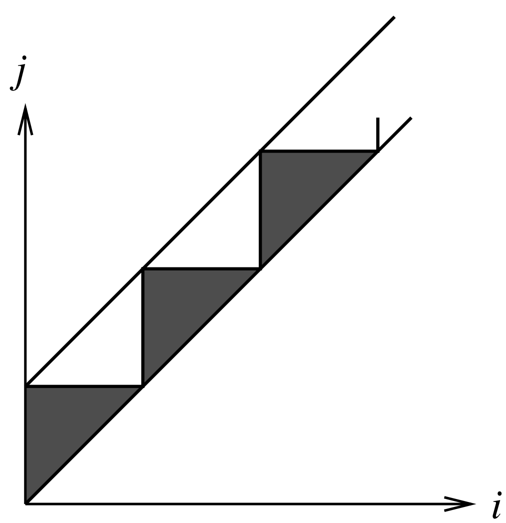

FIG 1. *A decomposition of Q_n.*

If $1 \le a < b \le m-1$, then $B_a^m < B_b^m$ and $W_a^m < W_b^m$. Consequently, (2.27) implies that

$$\text{(2.31)} \qquad \operatorname{Var} Q^{2,m}_{(n-1)m} \le 2(n-1)\left[\operatorname{Var} \Upsilon(B_1^m) + \operatorname{Var} \Upsilon(W_1^m)\right].$$

Because $\Upsilon(B_1^m)$ and $\Upsilon(W_1^m)$ are individually sums of not more than $\binom{m}{2}$-many ones,

$$\text{(2.32)} \qquad \operatorname{Var} Q^{2,m}_{(n-1)m} \le 2(n-1)m^2.$$

Let $Q_n^{(1)} := Q_n^{1,m}$ and $Q_n^{(2)} := Q_n^{2,m}$, where $m = m_n := n^{1/4}$ [say]. Then, $Q_n = Q_n^{(1)} + Q_n^{(2)}$, and (2.25) and (2.32) together imply that $\mathrm{E}Q^{(1)}_{(n-1)m} = o((n-1)m)$. Moreover, $\operatorname{Var} Q^{(2)}_{(n-1)m} = o((nm)^2)$. This gives us the desired decomposition of $Q_{(n-1)m}$. Now we complete the proof: Thanks to (2.22),

$$\text{(2.33)} \qquad \mathrm{E}Q^{(2)}_{(n-1)m} \sim nm \cdot \sum_{j=1}^{\infty} \mathrm{P}\{S_j = 0\} \quad \text{as } n \to \infty.$$

Therefore, the variance of $Q^{(2)}_{(n-1)m}$ is little-o of the square of its mean. This and the Chebyshev inequality together imply that $Q^{(2)}_{(n-1)m}/(nm)$ converges in probability to $\sum_{j=1}^{\infty} \mathrm{P}\{S_j = 0\}$. On the other hand, we know also that $Q^{(1)}_{(n-1)m}/(nm)$ converges to zero in $L^1(\mathrm{P})$ and hence in probability. Consequently, we can change variables and note that as $n \to \infty$,

$$\text{(2.34)} \qquad \frac{Q_{nm}}{nm} \to \sum_{j=1}^{\infty} \mathrm{P}\{S_j = 0\} \quad \text{in probability.}$$

If k is between $(n-1)m$ and nm, then

$$\text{(2.35)} \qquad \frac{Q_{(n-1)m}}{nm} \le \frac{Q_k}{k} \le \frac{Q_{nm}}{(n-1)m}.$$

This proves (2.21), and hence the theorem, in the transient case.

In order to derive the recurrent case, it suffices to prove that $Q_n/n \to \infty$ in probability as $n \to \infty$.

Let us choose and hold an integer $m \ge 1$—so that it does *not* grow with n—and observe that $Q_n \ge Q_n^{2,m}$ as long as n is sufficiently large. Evidently,

$$\text{(2.36)} \qquad \begin{aligned} \mathrm{E}Q_n^{2,m} &= \sum\sum_{\substack{1\le i<j\le n:\\ j<i+m}} \mathrm{P}\{S_j = S_i\} \\ &= (n-1)\sum_{k=1}^{m-1} \mathrm{P}\{S_k = 0\}. \end{aligned}$$

We may also observe that (2.32) continues to hold in the present recurrent setting. Together with the Chebyshev inequality, these computations imply that as $n \to \infty$,

$$\text{(2.37)} \qquad \frac{Q^{2,m}_{n(m-1)}}{n} \to \sum_{k=1}^{m-1} \mathrm{P}\{S_k = 0\} \quad \text{in probability.}$$

Because $Q_{n(m-1)} \geq Q^{2,m}_{n(m-1)}$, the preceding implies that

$$\lim_{n\to\infty} \mathrm{P}\left\{ \frac{Q_{n(m-1)}}{n} \geq \frac{1}{2}\sum_{k=1}^{m} \mathrm{P}\{S_k = 0\} \right\} = 1. \tag{2.38}$$

A monotonicity argument shows that $Q_{n(m-1)}$ can be replaced by Q_n without altering the end-result; see (2.35). By recurrence, if $\lambda > 0$ is any predescribed positive number, then we can choose [and fix] our integer m such that $\sum_{k=1}^{m} \mathrm{P}\{S_k = 0\} \geq 2\lambda$. This proves that $\lim_{n\to\infty} \mathrm{P}\{Q_n/n \geq \lambda\} = 1$ for all $\lambda > 0$, and hence follows the theorem in the recurrent case. $\square$

3. Proofs of the Main Results

Now we introduce a sequence $\{\xi_x\}_{x\in\mathbf{Z}^d}$ of random variables, independent [under P] of $\{q_i\}_{i=1}^{\infty}$ and the random walk $\{S_i\}_{i=0}^{\infty}$, such that

$$\mathrm{E}\xi_0 = 0, \quad \mathrm{E}\left(\xi_0^2\right) = \sigma^2, \quad \text{and} \quad \widehat{\mu}_3 := \mathrm{E}\left(|\xi_0|^3\right) < \infty. \tag{3.1}$$

Define

$$\widehat{h}^x_n := \left| \frac{L^x_n\left(L^x_n - 1\right)}{2} \right|^{1/2} \xi_x \qquad \text{for all } n \geq 1 \text{ and } x \in \mathbf{Z}^d. \tag{3.2}$$

Evidently, $\{\widehat{h}^x_n\}_{x\in\mathbf{Z}^d}$ is a sequence of [conditionally] independent random variables, under $\widehat{\mathrm{P}}$, and has the same [conditional] mean and variance as $\{h^x_n\}_{x\in\mathbf{Z}^d}$.

Lemma 3.1. *There exists a positive and finite constant $C_* = C_*(\sigma)$ such that if $f : \mathbf{R}^d \to \mathbf{R}$ is three time continuously differentiable, then for all $n \geq 1$,*

$$\left| \mathrm{E}f\left(\sum_{x\in\mathbf{Z}^d} \widehat{h}^x_n\right) - \mathrm{E}f(H_n) \right| \leq C_* M_f (\widehat{\mu}_3 + \mu_6) n \left| \sum_{j=0}^{n} \mathrm{P}\{S_j = 0\} \right|^2, \tag{3.3}$$

with $M_f := \sup_{x\in\mathbf{R}^d} |f'''(x)|$.

Proof. Temporarily choose and fix some $y \in \mathbf{Z}^d$, and notice that

$$\begin{aligned} &f(H_n) \\ &= f\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h^x_n\right) + f'\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h^x_n\right) h^y_n + \frac{1}{2} f''\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h^x_n\right) |h^y_n|^2 \\ &\quad + R_n, \end{aligned} \tag{3.4}$$

where $|R_n| \leq \frac{1}{6}\|f'''\|_\infty |h^y_n|^3$. It follows from this and Lemma 2.1 that

$$\begin{aligned} &\widehat{\mathrm{E}}f(H_n) \\ &= \widehat{\mathrm{E}}f\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h^x_n\right) + \frac{\sigma^2}{2} L^y_n\left(L^y_n - 1\right) \widehat{\mathrm{E}}f''\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h^x_n\right) + R^{(1)}_n, \end{aligned} \tag{3.5}$$

where

$$\begin{aligned} \left|R_n^{(1)}\right| &\le \frac{CM_f\mu_6}{12}\left|L_n^x\left(L_n^x-1\right)\right|^{3/2} \qquad \text{P-a.s.} \\ &\le \frac{CM_f\mu_6}{12}\left|L_n^y\right|^3. \end{aligned} \tag{3.6}$$

We proceed as in (3.4) and write

$$\begin{aligned} & f\left(\widehat{h}_n^y + \sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right) \\ &= f\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right) + f'\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right)\widehat{h}_n^y + \frac{1}{2}f''\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right)\left|\widehat{h}_n^y\right|^2 \\ &+ \widehat{R}_n, \end{aligned} \tag{3.7}$$

where $|\widehat{R}_n| \le \frac{1}{6}M_f|\widehat{h}_n^y|^3 \le \frac{1}{12\sqrt{2}}M_f|L_n^y|^3\,|\xi_y|^3$. It follows from this and Lemma 2.1 that

$$\begin{aligned} & \widehat{\mathrm{E}}f\left(\widehat{h}_n^y + \sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right) \\ &= \widehat{\mathrm{E}}f\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right) + \frac{\sigma^2}{2}L_n^y\left(L_n^y-1\right)\widehat{\mathrm{E}}f''\left(\sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right) + R_n^{(2)}, \end{aligned} \tag{3.8}$$

where $|R_n^{(2)}| \le \frac{1}{12\sqrt{2}}\widehat{\mu}_3 M_f|L_n^y|^3$. Define $C_* := (C+1)/2$ to deduce from the preceding and (3.5) that P-a.s.,

$$\left|\widehat{\mathrm{E}}f\left(\widehat{h}_n^y + \sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right) - \widehat{\mathrm{E}}f\left(\sum_{x\in\mathbf{Z}^d} h_n^x\right)\right| \le \frac{A}{6}|L_n^y|^3, \tag{3.9}$$

where $A := C_*M_f(\widehat{\mu}_3 + \mu_6)$. The preceding computes the effect of replacing the contribution of h_n^x to H_n by the independent quantity $\widehat{h}_n^y$, for each fixed y, and uses only the fact that the $\hat{h}$'s are a conditionally independent sequence with the same means and variances as their corresponding h's. Therefore, if we choose and fix another point $y \in \mathbf{Z}^d\setminus\{y\}$, then the very same constant A satisfies the following: Almost surely [P],

$$\left|\widehat{\mathrm{E}}f\left(\widehat{h}_n^z + \widehat{h}_n^y + \sum_{x\in\mathbf{Z}^d\setminus\{y,z\}} h_n^x\right) - \widehat{\mathrm{E}}f\left(\widehat{h}_n^y + \sum_{x\in\mathbf{Z}^d\setminus\{y\}} h_n^x\right)\right| \le \frac{A}{6}|L_n^z|^3. \tag{3.10}$$

And hence, the triangle inequality yields the following: P-a.s.,

$$\begin{aligned} & \left|\widehat{\mathrm{E}}f\left(\widehat{h}_n^z + \widehat{h}_n^y + \sum_{x\in\mathbf{Z}^d\setminus\{y,z\}} h_n^x\right) - \widehat{\mathrm{E}}f\left(\sum_{x\in\mathbf{Z}^d} h_n^x\right)\right| \\ &\qquad\qquad \le \frac{A}{6}\left(|L_n^y|^3 + |L_n^z|^3\right). \end{aligned} \tag{3.11}$$

Because $\sum_{x\in\mathbf{Z}^d} h_n^x = H_n$, it is now possible to see how we can iterate the previous inequality to find that P-a.s.,

$$\left|\widehat{\mathrm{E}}f\left(\sum_{x\in\mathbf{Z}^d}\widehat{h}_n^x\right) - \widehat{\mathrm{E}}f(H_n)\right| \le \frac{A}{6}\sum_{y\in\mathbf{Z}^d}|L_n^y|^3. \tag{3.12}$$

We take expectations and appeal to Lemma 2.2 to finish. □

Next, we prove Theorem 1.1.

Proof of Theorem 1.1. We choose, in Lemma 3.1, the collection $\{\xi_x\}_{x\in\mathbf{Z}^d}$ to be i.i.d. mean-zero normals with variance σ^2. Then, we apply Lemma 3.1 with $f(x) := g(x/n^{1/2})$ for a smooth bounded function g with bounded derivatives. This yields,

$$\left|\mathrm{E}g(H_n/n^{1/2}) - \mathrm{E}g\left(\frac{1}{n^{1/2}}\sum_{x\in\mathbf{Z}^d}\widehat{h}_n^x\right)\right| \le \frac{\mathrm{const}}{n^{1/2}}. \tag{3.13}$$

In this way,

$$\begin{aligned}\sum_{x\in\mathbf{Z}^d}\widehat{h}_n^x &\stackrel{\mathcal{D}}{=} \frac{\sigma}{\sqrt{2}}\left|\sum_{x\in\mathbf{Z}^d} L_n^x\,(L_n^x-1)\right|^{1/2} N(0\,,1) \qquad \text{under } \widehat{\mathrm{P}}\\ &= \frac{\sigma}{\sqrt{2}}\left|-n+\sum_{x\in\mathbf{Z}^d}(L_n^x)^2\right|^{1/2} N(0\,,1),\end{aligned} \tag{3.14}$$

where $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution, and $N(0\,,1)$ is a standard normal random variable under $\widehat{\mathrm{P}}$ as well as P. Therefore, in accord with Theorem 2.4,

$$\begin{aligned}\frac{1}{n^{1/2}}\sum_{x\in\mathbf{Z}^d}\widehat{h}_n^x &\stackrel{\mathcal{D}}{=} \frac{\sigma}{\sqrt{2}}\left|-1+\frac{1}{n}\sum_{x\in\mathbf{Z}^d}(L_n^x)^2\right|^{1/2} N(0\,,1)\\ &= o_{\widehat{\mathrm{P}}}(1) + \gamma^{1/2}\cdot N(0\,,\sigma^2),\end{aligned} \tag{3.15}$$

where $o_{\widehat{\mathrm{P}}}(1)$ is a term that converges to zero as $n\to\infty$ in $\widehat{\mathrm{P}}$-probability a.s. [P]. Equation (3.13) then completes the proof in the transient case. □

Theorem 1.2 relies on the following "coupled moderate deviation" result.

Proposition 3.2. *Suppose that S is recurrent. Consider a sequence $\{\epsilon_j\}_{j=1}^\infty$ of nonnegative numbers that satisfy the following:*

$$\lim_{n\to\infty}\epsilon_n^3 n\left|\sum_{k=1}^n \mathrm{P}\{S_k=0\}\right|^2 = 0. \tag{3.16}$$

Then for all compactly supported functions $f:\mathbf{R}^d\to\mathbf{R}$ that are infinitely differentiable,

$$\lim_{n\to\infty}\left|\mathrm{E}\left[f\left(\epsilon_n W_n\right)\right] - \mathrm{E}\left[f\left(\epsilon_n H_n\right)\right]\right| = 0. \tag{3.17}$$

Proof. We apply Lemma 3.1 with the ξ_x's having the same common distribution as q_1, and with $f(x) := g(\epsilon_n x)$ for a smooth and bounded function g with bounded derivatives. This yields,

$$\begin{aligned}&\left|\mathrm{E}\left[g\left(\epsilon_n \sum_{x\in\mathbf{Z}^d} \left|L_n^x\left(L_n^x-1\right)\right|^{1/2} Z(x)\right)\right] - \mathrm{E}\left[g\left(\epsilon_n H_n\right)\right]\right| \\ &\qquad \le 2C_* M_g \mu_6 n\epsilon_n^3 \left|\sum_{k=0}^n \mathrm{P}\{S_k=0\}\right|^2 \\ &\qquad = o(1),\end{aligned} \tag{3.18}$$

owing to Lemma (3.4).

According to Taylor's formula,

$$\begin{aligned}&g\left(\epsilon_n \sum_{x\in\mathbf{Z}^d} \left|L_n^x\left(L_n^x-1\right)\right|^{1/2} Z(x)\right) \\ &\quad = g\left(\epsilon_n \sum_{x\in\mathbf{Z}^d} Z(x)L_n^x\right) + \epsilon_n \sum_{x\in\mathbf{Z}^d}\left(\left|L_n^x\left(L_n^x-1\right)\right|^{1/2} - L_n^x\right) Z(x)\cdot R,\end{aligned} \tag{3.19}$$

where $|R| \le \sup_{x\in\mathbf{R}^d} |g'(x)|$. Thanks to (2.2), we can write the preceding as follows:

$$\begin{aligned}&g\left(\epsilon_n \sum_{x\in\mathbf{Z}^d} \left|L_n^x\left(L_n^x-1\right)\right|^{1/2} Z(x)\right) - g\left(\epsilon_n W_n\right) \\ &\qquad = \epsilon_n \sum_{x\in\mathbf{Z}^d}\left(\left|L_n^x\left(L_n^x-1\right)\right|^{1/2} - L_n^x\right) Z(x)\cdot R.\end{aligned} \tag{3.20}$$

Consequently, P-almost surely,

$$\begin{aligned}&\left|\widehat{\mathrm{E}}\left[g\left(\epsilon_n \sum_{x\in\mathbf{Z}^d} \left|L_n^x\left(L_n^x-1\right)\right|^{1/2} Z(x)\right)\right] - \widehat{\mathrm{E}}\left[g\left(\epsilon_n W_n\right)\right]\right| \\ &\quad \le \sup_{x\in\mathbf{R}^d} |g'(x)|\sigma\cdot\epsilon_n \left\{\widehat{\mathrm{E}}\left(\sum_{x\in\mathbf{Z}^d}\left(\left|L_n^x\left(L_n^x-1\right)\right|^{1/2} - L_n^x\right)^2\right)\right\}^{1/2}.\end{aligned} \tag{3.21}$$

We apply the elementary inequality $(a^{1/2}-b^{1/2})^2 \le |a-b|$—valid for all $a, b \ge 0$—to deduce that P-almost surely,

$$\begin{aligned}&\left|\widehat{\mathrm{E}}\left[g\left(\epsilon_n \sum_{x\in\mathbf{Z}^d} \left|L_n^x\left(L_n^x-1\right)\right|^{1/2} Z(x)\right)\right] - \widehat{\mathrm{E}}\left[g\left(\epsilon_n W_n\right)\right]\right| \\ &\qquad \le \sup_{x\in\mathbf{R}^d} |g'(x)|\sigma\cdot\epsilon_n \left\{\widehat{\mathrm{E}}\left(\sum_{x\in\mathbf{Z}^d} L_n^x\right)\right\}^{1/2} \\ &\qquad = \sup_{x\in\mathbf{R}^d} |g'(x)|\sigma\cdot\epsilon_n n^{1/2}.\end{aligned} \tag{3.22}$$

We take E-expectations and apply Lemma (3.4) to deduce from this and (3.18) that

$$|\mathrm{E}\left[g\left(\epsilon_n W_n\right)\right] - \mathrm{E}\left[g\left(\epsilon_n H_n\right)\right]| = o(1). \tag{3.23}$$

This completes the proof. ☐

Our proof of Theorem 1.2 hinges on two more basic lemmas. The first is an elementary lemma from integration theory.

Lemma 3.3. *Suppose $X := \{X_n\}_{n=1}^{\infty}$ and $Y := \{Y_n\}_{n=1}^{\infty}$ are $\mathbf{R}^d$-valued random variables such that: (i) X and Y each form a tight sequence; and (ii) for all bounded infinitely-differentiable functions $g : \mathbf{R}^d \to \mathbf{R}$,*

$$\lim_{n\to\infty} |\mathrm{E}g(X_n) - \mathrm{E}g(Y_n)| = 0. \tag{3.24}$$

Then, the preceding holds for all bounded continuous functions $g : \mathbf{R}^d \to \mathbf{R}$.

Proof. The proof uses standard arguments, but we repeat it for the sake of completeness.

Let $K_m := [-m\,, m]^d$, where m takes values in $\mathbf{N}$. Given a bounded continuous function $g : \mathbf{R}^d \to \mathbf{R}$, we can find a bounded infinitely-differentiable function $h_m : \mathbf{R}^d \to \mathbf{R}$ such that $|h_m - g| < 1/m$ on K_m. It follows that

$$\begin{aligned}|\mathrm{E}g(X_n) - \mathrm{E}g(Y_n)| &\le 2/m + |\mathrm{E}h_m(X_n) - \mathrm{E}h_m(Y_n)| \\ &\qquad + 2 \sup_{x\in\mathbf{R}^d} |g(x)| \big(\mathrm{P}\{X_n \not\in K_m\} + \mathrm{P}\{Y_n \not\in K_m\}\big).\end{aligned} \tag{3.25}$$

Consequently,

$$\begin{aligned}&\limsup_{n\to\infty} |\mathrm{E}g(X_n) - \mathrm{E}g(Y_n)| \\ &\qquad \le 2/m + 2 \sup_{x\in\mathbf{R}^d} |g(x)| \sup_{j\ge 1} \left(\mathrm{P}\{X_j \not\in K_m\} + \mathrm{P}\{Y_j \not\in K_m\}\right).\end{aligned} \tag{3.26}$$

Let m diverge and appeal to tightness to conclude that the left-had side vanishes. ☐

The final ingredient in the proof of Theorem 1.1 is the following harmonic-analytic result.

Lemma 3.4. *If $\epsilon_n := 1/a_n$, then (3.16) holds.*

Proof. Let ϕ denote the characteristic function of S_1. Our immediate goal is to prove that $|\phi(t)| < 1$ for all but a countable number of $t \in \mathbf{R}^d$. We present an argument, due to Firas Rassoul-Agha, that is simpler and more elegant than our original proof.

Suppose S_1' is an independent copy of S_1, and note that whenever $t \in \mathbf{R}^d$ is such that $|\phi(t)| = 1$, $D := \exp\{it\cdot(S_1 - S_1')\}$ has expectation one. Consequently, $\mathrm{E}(|D-1|^2) = \mathrm{E}(|D|^2) - 1 = 0$, whence $D = 1$ a.s. Because S_1 is assumed to have at least two possible values, $S_1 \neq S_1'$ with positive probability, and this proves that $t \in 2\pi\mathbf{Z}^d$. It follows readily from this that

$$\big\{t \in \mathbf{R}^d : |\phi(t)| = 1\big\} = 2\pi\mathbf{Z}^d, \tag{3.27}$$

and in particular, $|\phi(t)| < 1$ for almost all $t \in \mathbf{R}^d$.

By the inversion theorem (Spitzer [16] [**P3**(b), p. 57]), for all $n \geq 0$,

$$\mathrm{P}\{S_n = 0\} = \frac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} \{\phi(t)\}^n \, dt. \tag{3.28}$$

This and the dominated convergence theorem together tell us that $\mathrm{P}\{S_n = 0\} = o(1)$ as $n \to \infty$, whence it follows that

$$\sum_{k=1}^{n} \mathrm{P}\{S_k = 0\} = o(n) \quad \text{as } n \to \infty. \tag{3.29}$$

For our particular choice of ϵ_n we find that

$$\epsilon_n^3 n \left| \sum_{k=1}^{n} \mathrm{P}\{S_k = 0\} \right|^2 = \left(\frac{1}{n} \sum_{k=1}^{n} \mathrm{P}\{S_k = 0\} \right)^{1/2}, \tag{3.30}$$

and this quantity vanishes as $n \to \infty$ by (3.29). This proves the lemma. □

Proof of Theorem 1.2. Let $\epsilon_n := 1/a_n$. In light of Proposition 3.2, and Lemmas 3.3 and 3.4, it suffices to prove that the sequences $n \mapsto \epsilon_n W_n$ and $n \mapsto \epsilon_n H_n$ are tight.

Lemma 2.2, (2.2), and recurrence together imply that for all n large,

$$\begin{aligned} \mathrm{E}\left(|\epsilon_n W_n|^2\right) &= \sigma^2 \epsilon_n^2 \sum_{x \in \mathbf{Z}^d} \mathrm{E}\left(|L_n^x|^2\right) \\ &\leq \mathrm{const} \cdot \epsilon_n^2 n \sum_{k=1}^{n} \mathrm{P}\{S_k = 0\} \\ &= \mathrm{const}. \end{aligned} \tag{3.31}$$

Thus, $n \mapsto \epsilon_n W_n$ is bounded in $L^2(\mathrm{P})$, and hence is tight.

We conclude the proof by verifying that $n \mapsto \epsilon_n H_n$ is tight. Thanks to (2.4) and recurrence, for all n large,

$$\begin{aligned} \mathrm{E}\left(|\epsilon_n H_n|^2\right) &\leq \mathrm{const} \cdot \epsilon_n^2 \mathrm{E} \sum_{x \in \mathbf{Z}^d} (L_n^x)^2 \\ &\leq \mathrm{const} \cdot \epsilon_n^2 n \sum_{k=1}^{n} \mathrm{P}\{S_k = 0\} \\ &= \mathrm{const}. \end{aligned} \tag{3.32}$$

Confer with Lemma 2.2 for the penultimate line. Thus, $n \mapsto \epsilon_n H_n$ is bounded in $L^2(\mathrm{P})$ and hence is tight, as was announced. □

Acknowledgements

We wish to thank Siegfried Hörmann, Richard Nickl, Jon Peterson, and Firas Rassoul-Agha for many enjoyable discussions, particularly on the first portion of Lemma 3.4. We are grateful to Nadine Guillotin–Plantard for her valuable suggestions and typographical corrections, as well as providing us with references to her paper (Guillotine–Plantard [8]). Special thanks are extended to Firas Rassoul-Agha for providing us with his elegant argument that replaced our clumsier proof of the first part of Lemma 3.4. Finally, we thank two anonymous referees who made valuable suggestions and pointed out misprints.

References

[1] Bolthausen, E. (1989). A central limit theorem for two-dimensional random walks in random sceneries. *Ann. Probab.* **17** 108–115.

[2] Buffet, E. and Pulé, J. V. (1997). A model of continuous polymers with random charges. *J. Math. Phys.* **38** 5143–5152.

[3] Chen, X. (2008). Limit laws for the energy of a charged polymer. *Ann. Inst. H. Poincaré Probab. Statist.* **44** 638–672.

[4] Chung, K. L. and Fuchs, W. H. J. (1951). On the distribution of values of sums of random variables. *Mem. Amer. Math. Soc.* **6**.

[5] Derrida, B., Griffiths, R. B. and Higgs, R. G. (1992). A model of directed walks with random interactions. *Europhys. Lett.* **18** 361–366.

[6] Derrida, B. and Higgs, R. G. (1994). Low-temperature properties of directed walks with random self interactions. *J. Phys. A* **27** 5485–5493.

[7] Garel, T. and Orland, H. (1988). Mean-field model for protein folding. *Europhys. Lett.* **6** 307–310.

[8] Guillotin-Plantard, N. (2004). Sur la convergence faible des systèmes dynamiques échantillonnés. *Ann. Inst. Fourier* **54** 255–278.

[9] Kantor, Y. and Kardar, M. (1991). Polymers with self-interactions. *Europhys. Lett.* **14** 421–426.

[10] Kesten, H. and Spitzer, F. (1979). A limit theorem related to a new class of self-similar processes. *Z. Wahrsch. Verw. Gebiete* **50** 5–25.

[11] Liapounov, A. M. (1900). Sur une proposition de la théorie des probabilités. *Bull. de'l Acad. Impériale des Sci. St. Petérsbourg* **13** 359–386.

[12] Lindeberg, W. (1922). Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechtnung. *Math. Z.* **15** 211–225.

[13] Martìnez, S. and Petritis, D. (1996). Thermodynamics of a Brownian bridge polymer model in a random environment. *J. Phys. A* **29** 1267–1279.

[14] Obukhov, S. P. (1986). Configurational statistics of a disordered polymer chain. *J. Phys. A* **19** 3655–3664.

[15] Pólya, G. (1921). Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt in Straßennetz. *Math. Ann.* **84** 149–160.

[16] Spitzer, F. (1976). *Principles of Random Walks*, 2nd ed. Springer, New York–Heidelberg.

[17] Wittmer, J., Johner, A. and Joanny, J. F. (1993). Random and alternating polyampholytes. *Europhys. Lett.* **24** 263–268.