Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Conditional large deviations for a sequence of words

Matthias Birkner[1]

submitted: 22 November 2005

[1] Weierstraß-Institut
für Angewandte Analysis und Stochastik
Mohrenstraße 39
10117 Berlin
Germany

birkner@wias-berlin.de

No. 1072
Berlin 2005

2000 *Mathematics Subject Classification.* 60F10, 60G10, 82D60.

Key words and phrases. Conditional process level large deviations, quenched free energy.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Conditional large deviations for a sequence of words

Matthias Birkner

22 November 2005

Abstract

Cut an i.i.d. sequence (X_i) of "letters" into "words" according to an independent renewal process. Then one obtains an i.i.d. sequence of words, and thus the level three large deviation behaviour of this sequence of words is governed by the specific relative entropy. We consider the corresponding problem for the *conditional* empirical process of words, where one conditions on a typical underlying (X_i). We find that if the tails of the word lengths decay super-exponentially, the large deviations under the conditional distribution are again governed by the specific relative entropy, but the set of attainable limits is restricted.

We indicate potential applications of such a conditional LDP to the computation of the quenched free energy for directed polymer models with random disorder.

Key words: Conditional process level large deviations, quenched free energy.

MSC 2000 classification: 60F10, 60G10, 82D60.

1 Scenario and main result

Let E be a countable set ("letters" or "symbols"), $\nu \in \mathcal{P}(E)$ a probability measure on E with $\nu(x) > 0$ for all $x \in E$. Let $(X_i)_{i\in\mathbb{Z}}$ be an i.i.d.-ν sequence, $(\tau_j)_{j\in\mathbb{Z}}$ an independent i.i.d.-ρ sequence with values in $\{1, 2, \dots, \tau_{\max}\}$ resp. $\{1, 2, \dots\}$ if $\tau_{\max} = \infty$. In that case, we require that tail of the distribution of τ_0s decays super-exponentially in the following strong sense:

$$\exists\, C, \lambda, \epsilon > 0 \,:\, \forall n \,:\, \rho(\{n, n+1, \dots\}) \le C \exp(-\lambda\, n^{1+\epsilon}). \tag{1.1}$$

We also assume that the τs generate an aperiodic renewal process, i.e. $\gcd\{i : \rho_i > 0\} = 1$.

Cut out the X-sequence according to τ: Put $T_0 := 0$, $T_i := T_{i-1} + \tau_{i-1}$, $T_{-i} = T_{-i+1} - \tau_{-i}$ for $i > 0$,

$$Y^{(i)} = \big(X_{T_i}, X_{T_i+1}, \dots, X_{T_{i+1}-1}\big), \quad i \in \mathbb{Z} \tag{1.2}$$

with values in $\tilde{E} = \cup_{k=1}^{\infty} E^k$ ("words") (resp. $\tilde{E} = \cup_{k=1}^{\tau_{\max}} E^k$ if $\tau_{\max} < \infty$). By the independence properties of the ingredients, $Y = (Y^{(i)})_{i\in\mathbb{Z}}$ is then an i.i.d. sequence with marginal distribution

$$q^0\big((x_1, \dots, x_k)\big) := \mathbb{P}(Y^{(0)} = (x_1, \dots, x_k)) = \rho_k \prod_{i=1}^{k} \nu(x_i). \tag{1.3}$$

For a sequence $(Y^{(i)})$ with values in $\tilde{E}^{\mathbb{Z}}$ we write $L_i = |Y^{(i)}|$ for the "length" of the i-th word (in the present scenario, we have $L_i = \tau_i$, but it will be convenient to have a variable for word lengths also if Y does not arise from a construction with a τ-sequence). Note that we have a (left) shift $\theta : E^{\mathbb{Z}} \to E^{\mathbb{Z}}$ on letter sequences and a (left) shift $\tilde{\theta} : \tilde{E}^{\mathbb{Z}} \to \tilde{E}^{\mathbb{Z}}$ on word sequences. Let

$$R_N := \frac{1}{N} \sum_{i=0}^{N-1} \delta_{\tilde{\theta}^i Y} \tag{1.4}$$

be the empirical distribution process of the words with values in $\mathcal{P}(\tilde{E}^{\mathbb{Z}})$, the probability measures on sequences of words.

The sets E and $\tilde{E}$ are countable, so they are Polish spaces with the discrete metric. Then $E^{\mathbb{Z}}$ and $\tilde{E}^{\mathbb{Z}}$ are again metric spaces e.g. via $d_{A^{\otimes\mathbb{Z}}}\big((z_1, z_2, \dots), (z_1', z_2', \dots)\big) := \sum_{n-\infty}^{\infty} 2^{-|n|}\big(d_A(z_n, z_n') \wedge 1\big)$ for $A = E$ or $A = \tilde{E}$. This metric induces the product topology on $E^{\mathbb{Z}}$ resp. $\tilde{E}^{\mathbb{Z}}$. We equip $\mathcal{P}(\tilde{E}^{\mathbb{Z}})$ with the topology of weak convergence. Write $\mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})$ for the shift invariant probability measures on $\tilde{E}^{\mathbb{Z}}$, and $\mathcal{P}^{\mathrm{erg}}(\tilde{E}^{\mathbb{Z}})$ for the set of ($\tilde{\theta}$-shift) ergodic probability measures on $\tilde{E}^{\mathbb{Z}}$. Note that $\mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})$ is a closed subset of $\mathcal{P}(\tilde{E}^{\mathbb{Z}})$.

It is well known that the family of distributions $\mathscr{L}(R_N)$ satisfies a large deviation principle, the rate function is given by

$$H(Q; Q^0) = \lim_{N\to\infty} \frac{1}{N} h\big(Q|_{\mathscr{F}_N}; Q^0|_{\mathscr{F}_N}\big), \tag{1.5}$$

the specific relative entropy with respect to $Q^0 := \mathscr{L}(Y) = (q^0)^{\otimes\mathbb{Z}}$, see e.g. [3], Cor. 6.5.15 and Lemma 6.5.16. Here $\mathscr{F}_N = \sigma(Y_0, \dots, Y_{N-1})$, and $h(\mu; \mu')$ denotes the relative entropy of μ with respect to μ'. Our aim is to understand the almost sure large deviation behaviour of the family of random probability distributions

$$\mathscr{L}(R_N \,|\, X).$$

As $\mathcal{P}(E^{\mathbb{Z}})$ and $\mathcal{P}(\tilde{E}^{\mathbb{Z}})$ are Polish, we can and shall think in the following of a family of regular conditional distributions $\mathbb{P}(R_N \in \cdot\, | X)$.

Quantities involving the conditional expectation of exponential functionals of R_N appear naturally in the computation of the quenched free energy for polymer models in disordered media. In particular, the asymptotic evaluation of the free energy can be formulated as a conditional large deviation problem, and variational formulas as in Corollary 1 make the energy-entropy trade-off explicit. This potential application motivated our original interest in the question studied in this note, see Section 2 for more details.

It is natural to invert the cutting by concatenation: Let the concatenation operator $\kappa : \tilde{E}^{\mathbb{Z}} \to E^{\mathbb{Z}}$ be defined in the obvious way by

$$\kappa\big((\dots, y^{(-1)}, y^{(0)}, y^{(1)}, \dots)\big) = \big(\dots, y^{(-1)}_{\ell_{-1}}, y^{(0)}_1, y^{(0)}_2, \dots, y^{(0)}_{\ell_0}, y^{(1)}_1, y^{(1)}_2, \dots\big)$$

for $y^{(i)} = (y^{(i)}_1, \dots, y^{(i)}_{\ell_i}) \in \tilde{E}$. Note that $\kappa(y^{(\cdot)})$ has a "time origin": $\kappa(y^{(\cdot)})_0 = y^{(0)}_1$. One can imagine that because of the conditioning, which fixes a *typical* realisation of the X-sequence, the conditional law $\mathscr{L}(R_N \,|\, X)$ feels restrictions, and that some deviations, which are simply exponentially unlikely under the unconditional law, become actually impossible once a typical X is fixed. Let

$$\mathscr{R} := \Big\{ Q \in \mathcal{P}(\tilde{E}^{\mathbb{Z}}) : w - \lim_{L\to\infty} \frac{1}{L} \sum_{j=0}^{L-1} \delta_{\theta^j \kappa(Y)} = \nu^{\otimes\mathbb{Z}} \quad Q - \text{a.s.} \Big\}, \tag{1.6}$$

where $w - \lim$ denotes the limit with respect to the weak topology on $\mathcal{P}(E^{\mathbb{Z}})$. $Q \in \mathscr{R}$ means that under Q, the concatenation of words has almost surely the same asymptotic statistics as a typical realisation of (X_i). Obviously $Q^0 \in \mathscr{R}$. The following theorem is the main result of this note, it roughly states that under $\mathbb{P}(R_N \in \cdot | X)$, only such deviations can be realised which respect the restriction set $\mathscr{R}$.

Theorem 1. *Under Assumption (1.1), the following events occur with probability one:*

$$\limsup_N \frac{1}{N} \log \mathbb{P}\big(R_N \in F \,\big|\, X\big) \le - \inf_{Q \in F \cap \mathscr{R} \cap \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})} H(Q;Q^0) \quad \textit{for all closed } F \subset \mathcal{P}(\tilde{E}^{\mathbb{Z}}), \tag{1.7}$$

$$\liminf_N \frac{1}{N} \log \mathbb{P}\big(R_N \in G \,\big|\, X\big) \ge - \inf_{Q \in G \cap \mathscr{R} \cap \mathcal{P}^{\text{erg}}(\tilde{E}^{\mathbb{Z}})} H(Q;Q^0) \quad \textit{for all open } G \subset \mathcal{P}(\tilde{E}^{\mathbb{Z}}). \tag{1.8}$$

Corollary 1. *For any bounded continuous function $\Phi : \tilde{E}^{\mathbb{Z}} \to \mathbb{R}$ we have*

$$\begin{aligned} &\lim_N \frac{1}{N} \log \mathbb{E}\Big[\exp\big(N \int \Phi(y) R_N(dy)\big) \Big| X \Big] \\ &= \sup_{Q \in \mathscr{R} \cap \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})} \Big\{ \int \Phi(y) Q(dy) - H(Q;Q^0) \Big\} \quad \textit{a.s.} \end{aligned} \tag{1.9}$$

Remark 1. The same results hold for the "periodised" version

$$R_N^{\text{per}} := \frac{1}{N} \sum_{i=0}^{N-1} \delta_{\tilde{\theta}^i \left(Y^{(0)},\dots,Y^{(N-1)}\right)^{\text{per}}},$$

where $(y^{(0)}, \dots, y^{(N-1)})^{\text{per}}$ denotes the periodic extension of $(y^{(0)}, \dots, y^{(N-1)}) \in \tilde{E}^N$ to an element of $\tilde{E}^{\mathbb{Z}}$. The proofs are almost literally the same.

Theorem 1 is a full LDP for the family $\mathscr{L}(R_N|X)$, except for the restriction to *ergodic* Q in the lower bound. Removing this might require either a refinement of the conditional tilting employed in Section 5, or an argument that any $Q \in \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}}) \cap \mathscr{R}$ can be weakly approximated by shift ergodic Q_N which are also in $\mathscr{R}$, the latter requirement being the non-trivial part. On the other hand, the lower bound as stated in (1.8) suffices for the application in Corollary 1.

Remark 2. Theorem 1 does not hold in this form without assumptions on the tails of ρ. In fact, in a situation where ρ_n decays only algebraically, one can probe exponentially (in N, the number of pieces one wants to cut) far ahead into the X-sequence in order to find regions where X looks atypical.

For a concrete example, consider the following scenario: Let (X_i) be i.i.d. Ber(1/2), $\rho_n = C/n^a$, $a > 2$, so $m_\rho := \sum_n n\rho_n < \infty$. Put

$$\sigma_N := \min\{k \in \mathbb{N} : X_k = X_{k+1} = X_{k+[N(m_\rho+\epsilon)]} = 1\}.$$

Let $q^1(x_1, \dots, x_m) := \rho_m \mathbf{1}(x_1 = \dots = x_m = 1)$, and let $O \subset \mathcal{P}(\tilde{E}^{\otimes \mathbb{Z}})$ be a (small) neighbourhood of $(q^1)^{\otimes \mathbb{Z}}$. Under $(q^1)^{\otimes \mathbb{Z}}$, all words consist entirely of 1s. Note that $\log \sigma_N \sim N(m_\rho + \epsilon) \log 2$ by the Erdős-Rényi law and $\mathbb{P}(R_N \in O \,|\, X) \geq \rho_{\sigma_N}$ by (5.6) below, so

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}(R_N \in O|X) \geq \liminf_{N\to\infty} \frac{1}{N} \log \rho_{\sigma_N} > -\infty.$$

On the other hand, if (1.7) held true in this scenario, the answer would have to be $-\infty$, because $(q^1)^{\otimes \mathbb{Z}} \notin \mathscr{R}$.

By Lemma 5, (1.7) will hold with $\mathscr{R}$ replaced by $\overline{\mathscr{R}}$, but in view of Remark 4 in Section 3, this amounts essentially only to the unconditional upper bound, which we expect not to be sharp. The intuitive argument advocated on page 1, that any limiting Q must be built "on top" of a typical X-sequence, is not valid in general. In fact, when ρ has algebraic tails, there will be a trade-off on the exponential scale between how deep one probes into the fixed X-sequence, which allows to find more atypical regions, and the price for those long jumps. In view of the potential application to the computation of quenched free energies for polymer models in random media considered in Section 2, it appears a very interesting problem to find a quantitative description of this phenomenon. This question will be pursued in future work.

The rest of this paper is organised as follows: In Section 2 we indicate how Corollary 1, or rather, its analogue in a scenario where in contrast to Assumption (1.1), ρ has algebraic tails, could be used to represent the quenched free energy of directed polymer models with random disorder via a variational formula. We illustrate the use of Corollary 1 by expressing the quenched free energy of a modified polymer model. Coming back to the main plot, we give in Section 3 a useful characterisation of the property $Q \in \mathscr{R}$ under the additional constraint that Q has finite mean word lengths. This characterisation allows to make a connection between Q and an "underlying" i.i.d.-ν sequence. In Section 4, we basically prove the upper bound via comparison with the unconditional LDP, in Section 5 we prove the lower bound by a "conditional tilting" argument. The pieces are collected together in Section 6 to complete the proofs of Thm. 1 and Cor. 1.

2 Relation to quenched free energy computations

For models of directed polymers in disordered media, the quenched free energy involves quantities like

$$F_k(X) := \sum_{0 \leq j_1 < \dots < j_k} \prod_{i=1}^{k} \rho_{j_i - j_{i-1}} \exp\Big(\sum_{\ell=1}^{k} f\big((X_{j_{\ell-1}+1}, \dots, X_{j_\ell})\big)\Big) \tag{2.1}$$

for suitable functions $f : \tilde{E} \to \mathbb{R}$, where X is some i.i.d. sequence. By introducing an auxiliary sequence (τ_i) as in Section 1 and defining $(Y^{(i)})$ as in (1.2), this can be expressed as

$$\mathbb{E}\Big[\exp\Big(k\int f(y^{(\cdot)})\,r_k(dy^{(\cdot)})\Big)\Big|X\Big], \tag{2.2}$$

where $r_k := \pi_0 R_k$ is the image of R_k, as defined in (1.4), under the projection π_0 to the first coordinate, i.e. the empirical distribution of words. If the family $\mathscr{L}(R_k|X)$ almost surely satisfies a large deviation principle with a good, non-random rate function J, one can compute

$$\lim_{k\to\infty}\frac{1}{k}\log F_k(X) = \sup_{Q\in\mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})}\Big\{\int f(y)(\pi_0 Q)(dy) - J(Q)\Big\} \tag{2.3}$$

via Varadhan's Lemma (under suitable assumptions on f). Note that for such applications, indeed a level 2 large deviation principle would suffice, which one obtains by a contraction principle. On the other hand it seems that restrictions like (1.6) can only be expressed through level 3 objects.

As an example, let us consider the (modified) quenched free energy for the random heteropolymer model (see [1] and references there), defined as $\lim \frac{1}{N}\log Z^*_{N,X}$, where

$$Z^*_{N,X} = E\Big[\exp\big(\lambda\sum_{n=1}^{N}(X_n+h)\mathrm{sign}(S_n)\big); S_N = 0\Big],$$

where $\lambda, h \geq 0$, (S_n) is a symmetric simple random walk on $\mathbb{Z}$ starting at $S_0 = 0$, (X_n) are i.i.d. random variables, independent of S, taking the values ± 1 with probability 1/2 each, and E refers to expectation with respect to (S_n). In this context, if $S_n = 0$, "sign(S_n)" is defined as sign(S_{n-1}). We implicitly assume that N is even, otherwise $Z^*_{N,X} = 0$. This is a model for a polymer with a random composition of hydrophilic and hydrophobic monomers near an oil-water interface. The "letter" X_i models the affinity of monomer i towards different parts of the solvent. The free energy itself uses the same expression without the restriction on $\{S_N = 0\}$, this difference is irrelevant in the limit (see [1], Lemma 2).

Note that for the computation of the free energy, the details of the a priori measure on paths (S_n) are not important. All that matters is the fact that excursions from 0 are independent and symmetric, the only datum that is required to compute $Z^*_{N,X}$ is the distribution (ρ_n) of the excursion lengths: By decomposing the path $S_0, S_1, \ldots, S_N$ into excursions away from 0 and assigning independent random signs to the excursions, we can rewrite

$$Z^*_{N,X} = \sum_k \sum_{j_1<\cdots<j_k=N}\prod_{i=1}^{k}\rho_{j_i-j_{i-1}} \times \prod_{\ell=1}^{k}\cosh\Big(\lambda\sum_{i=j_{\ell-1}+1}^{j_\ell}(X_i+h)\Big), \tag{2.4}$$

where $\rho_n = P_0(S_1,\ldots,S_{n-1}\neq 0, S_n = 0)$ are the return probabilities for the random walk. Thus for $z \geq 0$ the (random) generating function of $Z^*_{N,X}$ is given by

$$\begin{aligned}
\theta(z) &= \sum_N z^N Z^*_{N,X} \\
&= \sum_N\sum_k\sum_{j_1<\cdots<j_k=N}\prod_{i=1}^{k}\rho_{j_i-j_{i-1}} \times \prod_{\ell=1}^{k}\Big\{z^{j_i-j_{i-1}}\cosh\Big(\lambda\sum_{i=j_{\ell-1}+1}^{j_\ell}(X_i+h)\Big)\Big\} \\
&= \sum_{k=1}^{\infty}F_k(X;z),
\end{aligned}$$

where $F_k(X;z)$ is the quantity defined in (2.1) with

$$f\big((x_1,\dots,x_\ell)\big) = \ell\log z + \log\cosh\Big(\lambda\sum_{i=1}^{\ell}(x_i+h)\Big). \tag{2.5}$$

Thus if we can (at least in principle) compute the almost sure asymptotic growth rate

$$\varphi(z) := \lim_{k\to\infty}\frac{1}{k}\log F_k(X;z)$$

via (2.3), we obtain that the radius of convergence of $\theta(z)$ is given by

$$r_\theta := \sup\{z\geq 0 : \varphi(z)<0\},$$

and hence the quenched free energy $f^{\mathrm{que}}(\lambda,h) = \lim N^{-1}\log Z^*_{N,X}$ can be represented as $f^{\mathrm{que}}(\lambda,h) = -\log r_\theta$. Note that the tails of ρ_n, the return probability of a 1-dimensional random walk, decay only algebraically in this scenario. In particular, ρ does not satisfy Assumption (1.1), so that the application of Corollary 1 to the computation of $\varphi(z)$ is not justified (and would, in view of Remark 2, almost certainly yield an incorrect result). We reiterate our statement from the end of Remark 2 that in view of the above considerations, it would be very interesting to extend Theorem 1 to the general case.

In order to illustrate the application of the conditional large deviation principle stated in Section 1, let us consider a modified model, where

the partition function $Z^*_{N,X}$ is given by (2.4) with ρ satisfying Assumption (1.1). (2.6)

This is a model for a situation where the polymer has a strong attraction towards the interface, as under the a priori measure excursions have short tails. So there can never be a de-pinning transition (as is the case for the original model, see [1]), but still for fixed realisation of (X_i), the polymer can try to optimise its configuration by grouping excursions according to stretches of X_is with the same sign, and there will be an energy-entropy trade-off. In this situation, the application of Corollary 1 is justified, and we can summarise the discussion above in the following

Proposition 1. *For the modified model (2.6) we have*

$$\varphi(z) = \sup_{Q\in\mathscr{R}\cap\mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})}\Big\{\int f(y)(\pi_0 Q)(dy) - H(Q;Q^0)\Big\} \quad a.s. \tag{2.7}$$

where in the notation of Section 1, $E=\{\pm 1\}$, $\nu(\pm 1)=1/2$, $q^0\big((x_1,\dots,x_\ell)\big) = 2^{-\ell}\rho_\ell$ for $(x_1,\dots,x_\ell)\in\{\pm1\}^\ell$, $Q^0=(q^0)^{\otimes\mathbb{Z}}$, f is defined in (2.5) and $\mathscr{R}$ in (1.6). The quenched free energy is given by

$$f^{\mathrm{que}}(\lambda,h) = -\log\Big(\sup\{z\geq 0:\varphi(z)<0\}\Big).$$

Note that (2.7) proves that $\varphi(z)$ is strictly increasing in z. Even though f is not bounded, we have for any $\gamma>1$

$$\sup_N\frac{1}{N}\log\mathbb{E}_{Q^0}\Big[\exp\Big(\gamma N\int f(y)\,(\pi_0 R_N)(dy)\Big)\Big] < \infty$$

by Assumption (1.1) and the fact that $f(y) \le C \times |y|$. Thus the application of Varadhan's Lemma is justified, cf e.g. Condition 4.3.3 in [3].

Proposition 1 allows comparison with the so called annealed free energy, defined as

$$f^{\mathrm{ann}}(\lambda, h) := \lim \frac{1}{N} \log \mathbb{E}\big[Z^*_{N,X}\big],$$

where $\mathbb{E}$ refers to expectation with respect to the distribution of (X_i). Defining $F_k^{\mathrm{ann}}(z) := \mathbb{E}[F_k(X;z)]$, $\varphi^{\mathrm{ann}}(z) := \lim k^{-1} \log F_k^{\mathrm{ann}}(z)$ we obtain

$$f^{\mathrm{ann}}(\lambda, h) = -\log\Big(\sup\{z \ge 0 : \varphi^{\mathrm{ann}}(z) < 0\}\Big)$$

in complete analogy with the reasoning above. As under the annealed measure the "marked" excursions $(Y^{(i)}))$ are i.i.d., $F_k^{\mathrm{ann}}(z) = (F_1^{\mathrm{ann}}(z))^k$ can be computed explicitly, however for the purpose of comparison with the quenched case, it is instructive to represent

$$\begin{aligned} \varphi^{\mathrm{ann}}(z) &= \log F_1^{\mathrm{ann}}(z) = \sup_{Q \in \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})} \Big\{ \int f(y)(\pi_0 Q)(dy) - H(Q;Q^0) \Big\} \\ &= \sup_{q \in \mathcal{P}(\tilde{E})} \Big\{ \int f(y) q(dy) - h(q;q^0) \Big\} = \log F_1^{\mathrm{ann}}(z) - \inf_{q \in \mathcal{P}(\tilde{E})} h(q; q^{*,\mathrm{ann}}), \end{aligned} \tag{2.8}$$

where $q^{*,\mathrm{ann}}\big((x_1 \dots, x_\ell)\big) = \frac{1}{F_1^{\mathrm{ann}}(z)} \rho_\ell \prod_{i=1}^{\ell} \nu(x_i) \times \exp f\big((x_1, \dots, x_\ell)\big)$ is (the marginal of) the unconstrained maximiser, which depends implicitly on z. Equality between the two sup-terms above stems from the fact that among all Q with given marginal $\pi_0 Q = q$, the specific relative entropy $H(Q;Q^0)$ is minimised by the product measure $Q = q^{\otimes \mathbb{Z}}$. From this together with Prop. 1 we see that in the model defined in (2.6), the "quenched to annealed bound" is always strict, i.e.

$$f^{\mathrm{que}}(\lambda, h) < f^{\mathrm{ann}}(\lambda, h) \quad \forall\, \lambda > 0, h \ge 0 \tag{2.9}$$

so there is no weak disorder regime.

In order to verify (2.9) suffices to check that $\varphi(z) < \varphi^{\mathrm{ann}}(z)$ for all $z \ge 0$. Fix $z \ge 0$, note that $Q^{*,\mathrm{ann}} := (q^{*,\mathrm{ann}})^{\otimes \mathbb{Z}} \notin \mathscr{R}$. A quick way to check this is as follows: In case $h > 0$, we see easily that $\sum_y y_1 q^{*,\mathrm{ann}}(y) > 0$, so $\lim_{L \to \infty} L^{-1} \sum_{j=0}^{L-1} \kappa(Y^{(\cdot)})_j > 0$ almost surely under $Q^{*,\mathrm{ann}}$, and hence $Q^{*,\mathrm{ann}} \notin \mathscr{R}$. On the other hand, if $h = 0$ we can observe that $\sum_{|y|=\ell} y_i y_j q^{*,\mathrm{ann}}(y) > 0$ for any $\ell \ge 2$, $1 \le i,j \le \ell$, i.e. letters are positively correlated under $q^{*,\mathrm{ann}}$, so $\lim_{L \to \infty} L^{-1} \sum_{j=0}^{L-1} \kappa(Y^{(\cdot)})_j \kappa(Y^{(\cdot)})_{j+1} > 0$ almost surely under $Q^{*,\mathrm{ann}}$, and hence again $Q^{*,\mathrm{ann}} \notin \mathscr{R}$.

As $\mathscr{R} \cap \mathscr{A}_M$ is compact by Lemma 3 and does not contain $Q^{*,\mathrm{ann}}$, we can find for any $M > 0$ a $\delta > 0$ such that $B_\delta(Q^{*,\mathrm{ann}}) \cap \mathscr{A}_M \subset \mathscr{R}^c$, and so

$$\varphi(z) \le \sup_{Q \in \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}}) \cap \big((B_\delta(Q^{*,\mathrm{ann}}))^c \cup \mathscr{A}_M^c\big)} \Big\{ \int f(y)(\pi_0 Q)(dy) - H(Q;Q^0) \Big\} < \varphi^{\mathrm{ann}}(z)$$

for a suitable choice of M and δ in view of (2.8).

3 A characterisation of the restriction set

When we cut the sequence X into pieces and then look at the empirical process of these pieces, we "loose the origin". In particular, the concatenation $\kappa(Y)$ under a limiting $Q \in \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})$ need not be shift invariant. For example, if we arrange the τs in such a way that the cut-points tend to occur before a certain pattern, then under R_N, the law of the concatenated sequence will have a (possibly atypical under $\nu^{\otimes\mathbb{Z}}$) inclination to begin with this pattern.

A way to reinstate shift-invariance (and in some way "get back the underlying i.i.d. sequence") which works when $\mathbb{E}_Q\, L_0 < \infty$ is to size-bias Q according to L_0 and then "randomise out the origin" – this is familiar from the theory of stationary renewal processes. Using this idea we obtain in this section a characterisation of the set $\mathscr{R}$ defined in (1.6).

For $Q \in \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})$ with $\mathbb{E}_Q\, L_0 < \infty$ let $\hat{Q} \in \mathcal{P}(\tilde{E}^{\mathbb{Z}})$ be defined by

$$\hat{Q}(Y^{(i)} = y^{(i)}, i = -k, \dots, k) = \frac{1}{\mathbb{E}_Q\, L_0} \mathbb{E}_Q\big[L_0 \mathbf{1}(Y^{(i)} = y^{(i)}, i = -k, \dots, k)\big] \tag{3.1}$$

(for any $k \in \mathbb{N}$, $y^{(i)} \in \tilde{E}$). Let $(\hat{Y}^{(i)})_{i\in\mathbb{Z}}$ have law $\hat{Q}$, given $\hat{Y}$, V uniform on $\{0, 1, \dots, L_0 - 1\}$, put

$$Z := \theta^V \kappa(\hat{Y}^{(\cdot)}).$$

We denote the distribution of Z obtained in this way by Ψ_Q to stress that it depends on Q.

We check that $\Psi_Q \in \mathcal{P}^{\mathrm{shift}}(E^{\mathbb{Z}})$: Fix $m \in \mathbb{N}$, $z_0, \dots, z_m \in E$. We have

$$\mathbb{P}(Z_0 = z_0, \dots, Z_m = z_m | \hat{Y}^{(\cdot)}) = \frac{1}{\hat{L}_0} \sum_{i=0}^{\hat{L}_0 - 1} \mathbf{1}(\kappa(\hat{Y}^{(\cdot)})_i = z_0, \dots, \kappa(\hat{Y}^{(\cdot)})_{i+m} = z_m),$$

hence

$$\begin{aligned}
&\mathbb{P}_{\Psi_Q}(Z_0 = z_0, \dots, Z_m = z_m) \\
&= \frac{1}{\mathbb{E}_Q\, L_0} \mathbb{E}_Q\big[L_0 \frac{1}{L_0} \sum_{i=0}^{L_0-1} \mathbf{1}(\kappa(Y^{(\cdot)})_i = z_0, \dots, \kappa(Y^{(\cdot)})_{i+m} = z_m)\big] \\
&= \frac{1}{\mathbb{E}_Q\, L_0} \mathbb{E}_Q\big[\sum_{i=0}^{L_0-1} \mathbf{1}(\kappa(Y^{(\cdot)})_i = z_0, \dots, \kappa(Y^{(\cdot)})_{i+m} = z_m)\big].
\end{aligned}$$

As Q is shift invariant,

$$\begin{aligned}
&\mathbb{E}_Q\big[\sum_{i=0}^{L_0-1} \mathbf{1}(\kappa(Y^{(\cdot)})_i = z_0, \dots, \kappa(Y^{(\cdot)})_{i+m} = z_m)\big] \\
&= \mathbb{E}_Q\big[\sum_{i=L_0+\dots+L_{k-1}}^{L_0+\dots+L_k-1} \mathbf{1}(\kappa(Y^{(\cdot)})_i = z_0, \dots, \kappa(Y^{(\cdot)})_{i+m} = z_m)\big]
\end{aligned}$$

for any $k \in \mathbb{Z}_+$, hence

$$\begin{aligned}
&\mathbb{P}_{\Psi_Q}(Z_0 = z_0, \dots, Z_m = z_m) \\
&= \frac{1}{M\mathbb{E}_Q\, L_0} \mathbb{E}_Q\big[\sum_{i=0}^{L_0+\dots+L_M-1} \mathbf{1}(\kappa(Y^{(\cdot)})_i = z_0, \dots, \kappa(Y^{(\cdot)})_{i+m} = z_m)\big]
\end{aligned}$$

for all $M \in \mathbb{N}$. Similarly, we have

$$\begin{aligned}&\mathbb{P}_{\Psi_Q}(Z_1 = z_0, \dots, Z_{m+1} = z_m)\\ &\quad = \frac{1}{M\mathbb{E}_Q L_0}\mathbb{E}_Q\Big[\sum_{i=0}^{L_0+\cdots+L_M-1} \mathbf{1}(\kappa(Y^{(\cdot)})_{i+1} = z_0, \dots, \kappa(Y^{(\cdot)})_{i+m+1} = z_m)\Big],\end{aligned}$$

consequently

$$\big|\mathbb{P}_{\Psi_Q}(Z_0 = z_0, \dots, Z_m = z_m) - \mathbb{P}_{\Psi_Q}(Z_1 = z_0, \dots, Z_{m+1} = z_m)\big| \le \frac{2}{M\mathbb{E}_Q L_0}.$$

Taking $M \to \infty$ we see that Ψ_Q is shift invariant.

Lemma 1. *Assume that $Q \in \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})$ satisfies $\mathbb{E}_Q L_0 < \infty$. Then we have $Q \in \mathscr{R}$ if and only if $\Psi_Q = \nu^{\otimes\mathbb{Z}}$. In this case, $\mathscr{L}_Q(\kappa(Y)) \ll \nu^{\otimes\mathbb{Z}}$.*

Proof. Let $\Psi_Q = \nu^{\otimes\mathbb{Z}}$. Then under $\hat{Q}$, the sequence $\kappa(Y)$ almost surely has the 'right' asymptotic pattern frequencies (i.e. $\lim N^{-1}\sum_{i=0}^{N-1} f_z(\theta^i\kappa(Y)) = \nu(z)$ for any $z \in \tilde{E}$, where $f_z(x) = \mathbf{1}(x_1 = z_1, \dots x_\ell = z_\ell)$). As $Q \ll \hat{Q}$ (in fact, the density is $(E_Q L_0)/L_0$, which is > 0, bounded), the same holds true for Q, i.e. $Q \in \mathscr{R}$.

Now assume that $Q \in \mathscr{R}$. As $\hat{Q} \ll Q$, the sequence Z_i, $i \in \mathbb{Z}$ under Ψ_Q also has the 'right' asymptotic pattern frequencies, i.e.

$$\lim_N \frac{1}{N}\sum_{i=0}^{N-1} \mathbf{1}(Z_i = z_0, \dots, Z_{i+m} = z_m) = \prod_{j=0}^{m} \nu(z_m) \quad \text{almost surely} \tag{3.2}$$

for any $m \in \mathbb{N}$, $z_0, \dots, z_m \in E$. It suffices to verify that any shift invariant sequence (Z_i) satisfying (3.2) is in fact an i.i.d.-ν sequence. The limit on the left-hand side of (3.2) is equal to

$$\mathbb{P}\big(Z_0 = z_0, \dots, Z_m = z_m \big| \mathscr{I}\big),$$

where $\mathscr{I}$ is the shift-invariant σ-field. Thus

$$\mathbb{P}(Z_0 = z_0, \dots, Z_m = z_m) = \mathbb{E}\Big[\mathbb{P}\big(Z_0 = z_0, \dots, Z_m = z_m \big| \mathscr{I}\big)\Big] = \prod_{j=0}^{m} \nu(z_m),$$

so that indeed $\mathscr{L}(Z) = \nu^{\otimes\mathbb{Z}}$.

Now assume that $\Psi_Q = \nu^{\otimes\mathbb{Z}}$ and let $\mathcal{A} \subset E^{\mathbb{Z}}$ be a (measurable) $\nu^{\otimes\mathbb{Z}}$-null set. Then we have

$$0 = \nu^{\otimes\mathbb{Z}}(\mathcal{A}) = \Psi_Q(\mathcal{A}) = \frac{1}{\mathbb{E}_Q L_0}\mathbb{E}_Q\Big[\sum_{i=0}^{L_0-1} \mathbf{1}_{\mathcal{A}}(\theta^i\kappa(Y))\Big],$$

so in particular $Q(\kappa(Y) \in \mathcal{A}) = 0$. This proves that $\mathscr{L}_Q(\kappa(Y)) \ll \nu^{\otimes\mathbb{Z}}$. □

Remark 3. If $Q \in \mathscr{R}$ and $\mathbb{E}_Q L_0 < \infty$, by the above there is a random V such that under $\hat{Q}$, $\theta^V\kappa(Y)$ is distributed like an i.i.d.-ν sequence. We can "invert" this relation: There is (on some probability space) a random pair (Δ, Z) with values in $\mathbb{Z} \times E^{\mathbb{Z}}$ such that $\mathscr{L}(Z) = \nu^{\otimes\mathbb{Z}}$ and $\mathscr{L}(\theta^\Delta Z) = \mathscr{L}_{\hat{Q}}(\kappa(Y))$. For example, one can take (Y, V) under $\hat{Q}$, then define $Z := \theta^V\kappa(Y)$, $\Delta := -V$.

Note that the mappings $Q \mapsto \hat{Q}$, $Q \mapsto \Psi_Q$ are not continuous with respect to the weak topology on $\mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})$ (because $\tilde{E}^{\mathbb{Z}} \ni (y^{(i)})_i \mapsto |y^{(0)}|$ is not bounded, so weak convergence need not imply convergence of the first moment of piece lengths). On the other hand, we have

Lemma 2. *Assume that $Q_N \in \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})$ converge weakly to Q_∞ and that additionally $\mathbb{E}_{Q_N}[L_0] \to \mathbb{E}_{Q_\infty}[L_0]$ as $N \to \infty$. Then*

$$\hat{Q}_N \to \hat{Q}_\infty \text{ weakly on } \mathcal{P}(\tilde{E}^{\mathbb{Z}}) \quad \text{and} \quad \Psi_{Q_N} \to \Psi_{Q_\infty} \text{ weakly on } \mathcal{P}(E^{\mathbb{Z}}).$$

Proof. Note that by the assumptions, the family $\{\mathscr{L}_{Q_N}(L_0), N \in \mathbb{N}\}$ is uniformly integrable. Hence also for any $k \in \mathbb{N}$, $y^{(i)} \in \tilde{E}$, the family $\{\mathscr{L}_{Q_N}(L_0 \mathbf{1}(Y^{(i)} = y^{(i)}, i = -k, \dots, k)), N \in \mathbb{N}\}$ is uniformly integrable. This implies

$$\hat{Q}_N(Y^{(i)} = y^{(i)}, i = -k, \dots, k) \to \hat{Q}_\infty(Y^{(i)} = y^{(i)}, i = -k, \dots, k).$$

Similarly, because $0 \le \sum_{i=0}^{L_0 - 1} \mathbf{1}(\kappa(Y^{(\cdot)})_i = z_0, \dots, \kappa(Y^{(\cdot)})_{i+m} = z_m) \le L_0$ (for any $m \in \mathbb{N}$, $z_i \in E$), we conclude that

$$\Psi_{Q_N}(Z_0 = z_0, \dots, Z_m = z_m) \to \Psi_{Q_\infty}(Z_0 = z_0, \dots, Z_m = z_m).$$

□

Lemma 3. *Let*

$$\mathscr{A}_M := \{Q \in \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}}) : H(Q; Q^0) \le M\}, \quad M \ge 0$$

be the level sets of the rate function $Q \mapsto H(Q; Q^0)$. For any M, the set $\mathscr{R} \cap \mathscr{A}_M$ is closed (in the weak topology on $\mathcal{P}(\tilde{E}^{\mathbb{Z}})$). In case $\tau_{\max} < \infty$, $\mathscr{R}$ is closed.

Proof. Let $(Q_N) \subset \mathscr{R} \cap \mathscr{A}_M$, assume $Q_N \to_{N \to \infty} Q_\infty$ weakly. As $h(\mathscr{L}_{Q_N}(L_0); \rho) \le H(Q_N; Q^0) \le M$ for all N, where $h(\mathscr{L}_{Q_N}(L_0); \rho)$ is the relative entropy of the length of the first word under Q_N with respect to ρ, we obtain

$$Q_N(\{L_0 \ge n\}) \le \frac{\log 2 + h(\mathscr{L}_{Q_N}(L_0); \rho)}{\log(1 + 1/\rho(\{n, n+1, \dots\}))} \le \frac{C' M}{n^{1+\epsilon}}. \tag{3.3}$$

by (1.1) and the entropy inequality (see e.g. [5], Prop. 8.2 in Appendix 1). This uniform bound on the tails of the word lengths under Q_N implies $\mathbb{E}_{Q_N} L_0 \to \mathbb{E}_{Q_\infty} L_0$, and thus $Q_\infty \in \mathscr{R}$ by Lemma 2 and Lemma 1. $Q_\infty \in \mathscr{A}_M$ because the level sets are compact.

In the situation $\tau_{\max} < \infty$ we have $\mathbb{E}_{Q_N} L_0 \to \mathbb{E}_{Q_\infty} L_0$ automatically, and the rest of the argument remains unchanged. □

Remark 4. In case $\tau_{\max} = \infty$, the set $\mathscr{R}$ is not closed in the weak topology. In fact,

$$\overline{\mathscr{R}} \supset \{Q \in \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}}) : \mathbb{E}_Q[L_0] < \infty\}.$$

Proof. Fix an arbitrary Q in $\mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})$ satisfying $\mathbb{E}_Q L_0 < \infty$. Let $\tilde{q} \in \mathcal{P}(\tilde{E})$ be given by

$$\tilde{q}\big((x_1, \dots, x_n)\big) = \frac{C}{n^{-3/2}} \prod_{i=1}^{n} \nu(x_i),$$

i.e. the length of the word has heavy tails, given the length is n, it looks like n independent draws from ν. Define Q_N as follows: under $\tilde{Q}_N$, the blocks $(Y^{(kN)}, Y^{(kN+1)}, \dots, Y^{((k+1)N-1)})$, $k \in \mathbb{Z}$, are i.i.d, $\mathscr{L}_{\tilde{Q}_N}((Y^{(0)}, \dots, Y^{(N-1)})) = \tilde{q} \otimes Q|_{\sigma(Y^{(1)}, \dots, Y^{(N-1)})}$. Q_N is defined as $\tilde{Q}_N$ with randomised origin, formally $Q_N = N^{-1} \sum_{i=0}^{N-1} \tilde{Q}_N \circ \tilde{\theta}^i$. Then we have $Q_N \in \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})$ (in fact even $Q_N \in \mathcal{P}^{\text{erg}}(\tilde{E}^{\mathbb{Z}})$), $Q_N \to Q$ weakly. Finally, each $Q_N \in \mathscr{R}$ because the word length under $\tilde{q}$ has no mean: imagine pointing at position U in $\kappa(Y)$ under Q_N, where $U \sim \text{Unif}(\{1, \dots, L\})$. As $L \to \infty$, the probability tends to one that one actually looks inside a "$\tilde{q}$-word" of the concatenation, where the pattern frequencies are what they ought to be in a $\nu^{\otimes \mathbb{Z}}$-sequence. □

4 Conditioning and the restriction set, upper bound

First we observe that an unconditional upper bound is automatically also an upper bound for the conditional distributions:

Lemma 4. *For any closed $F \subset \mathcal{P}(\tilde{E}^{\mathbb{Z}})$ we have*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(R_N \in F \,|\, X) \le - \inf_{Q \in F \cap \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})} H(Q; Q^0) \qquad a.s. \tag{4.1}$$

Proof. Write $I(F) := \inf_{Q \in F \cap \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})} H(Q; Q^0)$. For $\epsilon > 0$ we have by the unconditional LDP

$$\begin{aligned}
&\mathbb{P}\big(\mathbb{P}(R_N \in F \,|\, X) \ge \exp(-N(I(F) - 2\epsilon)\big) \\
&\le \; e^{N(I(F)-2\epsilon)} \mathbb{E}\big[\mathbb{P}(R_N \in F \,|\, X)\big] = e^{N(I(F)-2\epsilon)} \mathbb{P}(R_N \in F) \\
&\le \; e^{N(I(F)-2\epsilon)} e^{-N(I(F)-\epsilon)} = e^{-\epsilon N}
\end{aligned}$$

for N large enough, and hence

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(R_N \in F \,|\, X) \le -I(F) - 2\epsilon \qquad \text{a.s.}$$

by the Borel-Cantelli Lemma. Take $\epsilon \to 0$ to conclude. □

Lemma 5. *For any closed $F \subset \mathcal{P}(\tilde{E}^{\mathbb{Z}})$ and $M \ge 0$ we have a.s.*

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(R_N \in F \,|\, X) \le \Big(- \inf_{Q \in F \cap \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}}) \cap \mathscr{A}_M \cap \mathscr{R}} H(Q; Q^0) \Big) \vee \big(- M \big). \tag{4.2}$$

Proof. First note that even though R_N is not exactly shift-invariant because of boundary terms, it is nearly so: for any weak neighbourhood O of $\mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})$, there is n_0 such that $R_N \in O$ for $N \ge n_0$. As $\mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})$ is closed in the weak topology, we can restrict to $F \cap \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})$ in the right-hand side of (4.2).

Fix $\delta > 0$. As $Q \mapsto H(Q; Q^0)$ is a good rate function, the level sets $\mathscr{A}_M$ are compact. Thus there exist $n_0 \in \mathbb{N}$ and $Q_1, \dots, Q_{n_0} \in \mathscr{A}_M$ such that $\mathscr{A}_M \subset \cup_{i=1}^{n_0} B_\delta(Q_i)$, where $B_\delta(Q)$ denotes the open ball of radius δ around $Q \in \mathcal{P}(\tilde{E}^{\mathbb{Z}})$. Let

$$\begin{aligned}
J_1 &:= \{1 \le i \le n_0 : B_\delta(Q_i) \cap F \ne \emptyset\}, \\
J_2 &:= \{j \in J_1 : B_\delta(Q_j) \cap \mathscr{R} \ne \emptyset\}.
\end{aligned}$$

Note that

$$\begin{aligned}
\mathbb{P}(R_N \in F \,|\, X) \;&\le\; \mathbb{P}\Big(R_N \in \cup_{i\in J_1} B_\delta(Q_i) \,\Big|\, X\Big) \\
&\quad + \mathbb{P}\Big(R_N \in \big(\cup_{i=1}^n B_\delta(Q_i)\big)^c \,\Big|\, X\Big) \\
&\le\; \mathbb{P}\Big(R_N \in \cup_{i\in J_2} B_\delta(Q_i) \,\Big|\, X\Big) + \sum_{j\in J_1\setminus J_2} \mathbb{P}\Big(R_N \in B_\delta(Q_j) \,\Big|\, X\Big) \\
&\quad + \mathbb{P}\Big(R_N \in \big(\cup_{i=1}^n B_\delta(Q_i)\big)^c \,\Big|\, X\Big) \\
&\le\; \mathbb{P}\Big(R_N \in \big(F \cap \mathscr{A}_M \cap \mathscr{R}\big)_{2\delta} \,\Big|\, X\Big) + \sum_{j\in J_1\setminus J_2} \mathbb{P}\Big(R_N \in B_\delta(Q_j) \,\Big|\, X\Big) \\
&\quad + \mathbb{P}\Big(R_N \in \big(\cup_{i=1}^n B_\delta(Q_i)\big)^c \,\Big|\, X\Big),
\end{aligned}$$

where $(\mathcal{B})_\epsilon := \{Q : d(Q,\mathcal{B}) < \epsilon\}$.

We have $\mathbb{P}(R_N \in \mathscr{R} \,|\, X) = 1$ by ergodicity of (X_i), hence $\mathbb{P}(R_N \in B_\delta(Q_j) \,|\, X) = 0$ for $j \in J_2 \setminus J_1$. Furthermore, $\big(\cup_{i=1}^n B_\delta(Q_i)\big)^c$ is closed and contained in $\mathscr{A}_M^c$, so

$$\inf_{Q\in\big(\cup_{i=1}^n B_\delta(Q_i)\big)^c} H(Q;Q^0) \ge M,$$

and hence

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big(R_N \in \big(\cup_{i=1}^n B_\delta(Q_i)\big)^c \,\Big|\, X\Big) \le -M \qquad \text{a.s.}$$

by Lemma 4. Again by Lemma 4, we have a.s.

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big(R_N \in \big(F\cap\mathscr{A}_M\cap\mathscr{R}\big)_{2\delta} \,\Big|\, X\Big) \le -\inf\Big\{H(Q;Q^0) : Q \in \overline{\big(F\cap\mathscr{A}_M\cap\mathscr{R}\cap\mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})\big)_{2\delta}}\Big\},$$

so

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big(R_N \in F \,\Big|\, X\Big) \le \Big(-\inf_{Q\in\overline{\big(F\cap\mathscr{A}_M\cap\mathscr{R}\cap\mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})\big)_{2\delta}}} H(Q;Q^0)\Big) \vee \big(-M\big)$$

Taking $\delta \to 0$ and observing that $\mathscr{A}_M \cap \mathscr{R}$ is closed by Lemma 3 yields the claim. □

5 Lower bound

For a random variable W and a probability measure Q, we will in the following be writing $Q(W)$ for the random variable $\phi(W)$, where $\phi(w) = Q(W = w)$. $Q(W_1,\dots,W_N)$, etc is defined analogously.

Lemma 6. *Assume that $(Z_i)_{i\in\mathbb{N}}$ with values in $E^{\mathbb{Z}}$ has distribution P' and $P' \ll P = \nu^{\otimes\mathbb{Z}}$. Then we have*

$$\lim_{N\to\infty} -\frac{1}{N}\log P'(Z_1,\dots,Z_N) = -\sum_x \nu(x)\log\nu(x) \quad P'\text{-a.s.}$$

(i.e. the specific entropy of P' equals that of P).

Proof. Let $dP' = \Phi dP$, where $\Phi = \phi(Z_1, Z_2, \dots)$, $\mathscr{F}_n := \sigma(Z_1, \dots, Z_n)$. Then

$$\begin{aligned} P'(z_1, \dots, z_N) &= \mathbb{E}_P[\phi(Z)\mathbf{1}(Z_1 = z_1, \dots, Z_n = z_n)] \\ &= \nu(z_1) \cdots \nu(z_N) \times \int \nu^{\otimes \mathbb{N}}(dz')\phi(z_1, \dots, z_n, z'_1, z'_2, \dots), \end{aligned}$$

so

$$P'(Z_1, \dots, Z_n) = \nu(Z_1) \cdots \nu(Z_N)\mathbb{E}_P[\Phi|\mathscr{F}_n].$$

Now $\mathbb{E}_P[\Phi|\mathscr{F}_n] \to \Phi$ P- (and hence also P'-a.s.), thus

$$\frac{1}{N} \log P'(Z_1, \dots, Z_n) = \frac{1}{N} \sum_{i=1}^{N} \log \nu(Z_i) + O(1/N) \to \sum_x \nu(x) \log \nu(x).$$

□

For an $\tilde{E}^{\mathbb{Z}}$-valued random variable $(Y^{(i)})_{i \in \mathbb{Z}}$, $m \le n$, let us write $Z_{m,n} := \kappa(Y^{(\cdot)})|_{m \dots n}$ for the restriction of $\kappa(Y^{(\cdot)})$ to $\{m, m+1, \dots, n\}$.

Lemma 7. *Assume $Q \in \mathscr{R} \cap \mathcal{P}^{\text{shift}}(\tilde{E}^{\mathbb{Z}})$ is shift-ergodic and satisfies $m_Q := \mathbb{E}_Q L_0 < \infty$. Then for any $\epsilon > 0$ there is $\delta(\epsilon)$ such that*

$$\begin{aligned} H(Q; Q^0) - \delta(\epsilon) &\le \liminf_{N \to \infty} \frac{1}{N} \log \left(\frac{Q(L_0, \dots, L_{N-1}|Z_{0,N(m_Q+\epsilon)})}{\prod_{i=0}^{N-1} \rho_{L_i}} \right) \\ &\le \limsup_{N \to \infty} \frac{1}{N} \log \left(\frac{Q(L_0, \dots, L_{N-1}|Z_{0,N(m_Q+\epsilon)})}{\prod_{i=0}^{N-1} \rho_{L_i}} \right) \le H(Q; Q^0) + \delta(\epsilon), \end{aligned}$$

holds Q-almost surely, and $\delta(\epsilon) \to 0$ as $\epsilon \to 0$.

Note that the term appearing in lim inf and lim sup is approximately the specific relative entropy of the word lengths given the concatenation.

Proof. Note that

$$Q(L_0, \dots, L_{N-1}|Z_{0,N(m_Q+\epsilon)}) = \frac{Q(L_0, \dots, L_{N-1}, Z_{0,N(m_Q+\epsilon)})}{Q(Z_{0,N(m_Q+\epsilon)})}.$$

By Lemma 1 and Lemma 6, we have

$$\lim_{N \to \infty} -\frac{1}{N} \log Q(Z_{0,N(m_Q+\epsilon)}) = -(m_Q + \epsilon) \sum_{x \in E} \nu(x) \log \nu(x) \qquad Q\text{-a.s.} \tag{5.1}$$

Furthermore,

$$\lim_{N \to \infty} \frac{1}{N} \log \prod_{i=0}^{N-1} \rho_{L_i} = \mathbb{E}_Q \log \rho_{L_0} \qquad Q\text{-a.s.} \tag{5.2}$$

by ergodicity of Q. Fix ϵ' such that $(1 + \epsilon')m_Q > m_Q + \epsilon$, denote

$$B_N := \{L_0 + \dots + L_{N-1} \le N(m_Q + \epsilon) \le L_0 + \dots + L_{N(1+\epsilon')}\}.$$

Observe that on B_N,

$$Q(Y_0, Y_1, \dots, Y_{N(1+\epsilon')}) \le Q(L_0, \dots, L_{N-1}, Z_{0,N(m_Q+\epsilon)}) \le Q(Y_0, Y_1, \dots, Y_{N-1}).$$

As $Q(\cup_M \cap_{N \ge M} B_N) = 1$ by ergodicity, we obtain

$$\begin{aligned}
\limsup_{N\to\infty} \frac{1}{N} \log Q(L_0, \dots, L_{N-1}, Z_{0,N(m_Q+\epsilon)}) &\le -H(Q), \\
\liminf_{N\to\infty} \frac{1}{N} \log Q(L_0, \dots, L_{N-1}, Z_{0,N(m_Q+\epsilon)}) &\ge -(1+\epsilon')H(Q),
\end{aligned}$$

where

$$H(Q) = \lim -\frac{1}{N} \log Q(Y_0, Y_1, \dots, Y_{N-1})$$

is the specific entropy of Q. This together with (5.1) and (5.2) shows that lim inf and lim sup appearing in the statement of the lemma are given by

$$-m_Q \sum_{x\in E} \nu(x) \log \nu(x) - \mathbb{E}_Q \log \rho_{L_0} - H(Q) \tag{5.3}$$

plus terms that are $O(\epsilon)$. It remains to verify that for Q satisfying the assumptions of the lemma, $H(Q; Q^0)$ is given by (5.3): Let $\mathscr{F}_N := \sigma(Y_0, \dots, Y_{N-1})$. By the form of Q^0,

$$\begin{aligned}
&\frac{1}{N} h(Q|_{\mathcal{F}_N}; Q^0|_{\mathcal{F}_N}) \\
&= \frac{1}{N} \sum_{y^{(0)},\dots,y^{(N-1)}} Q(y^{(0)}, \dots, y^{(N-1)}) \log \frac{Q(y^{(0)}, \dots, y^{(N-1)})}{Q^0(y^{(0)}, \dots, y^{(N-1)})} \\
&= \frac{1}{N} \sum_{y^{(0)},\dots,y^{(N-1)}} Q(y^{(0)}, \dots, y^{(N-1)}) \log Q(y^{(0)}, \dots, y^{(N-1)}) \\
&\quad - \frac{1}{N} \sum_{y^{(0)},\dots,y^{(N-1)}} Q(y^{(0)}, \dots, y^{(N-1)}) \sum_{i=0}^{N-1} \log \rho_{|y^{(i)}|} \\
&\quad - \frac{1}{N} \sum_{y^{(0)},\dots,y^{(N-1)}} Q(y^{(0)}, \dots, y^{(N-1)}) \sum_{j=0}^{|y^{(0)}|+\dots+|y^{(i)}|-1} \log \nu(\kappa(y^{(0)}, \dots, y^{(N-1)})_j) \\
&\to -H(Q) - \mathbb{E}_Q \log \rho_{L_0} - m_Q \sum_x \nu(x) \log \nu(x)
\end{aligned}$$

by ergodicity of Q and Lemma 6. □

Lemma 8. *Let $Q \in \mathscr{R} \cap \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})$ be shift-ergodic and satisfy $m_Q := \mathbb{E}_Q L_0 < \infty$, let $O \subset \mathcal{P}(\tilde{E}^{\mathbb{Z}})$ be an open neighbourhood of Q. Then we have*

$$\liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}(R_N \in O | X) \ge -H(Q; Q^0) \qquad a.s. \tag{5.4}$$

Proof. Let $\hat{Q}$ be defined from Q via (3.1) as in Section 3. By Remark 3, there is a measurable function $f : E^{\mathbb{Z}} \times [0,1] \to \mathbb{Z}$ such that if $U \sim$ uniform$([0,1])$ independent of X, and $\tilde{\Delta} := f(X, U)$, then $\mathscr{L}(\theta^{\tilde{\Delta}}(X)) = \mathscr{L}_{\hat{Q}}(\kappa(Y))$ (in fact, we can assume $\tilde{\Delta} \le 0$). Let us further assume that we have set up things on a joint probability space in such a way that $Y^{(\cdot)} \sim \hat{Q}$ and

$$\kappa(Y^{(\cdot)}) = \theta^{\tilde{\Delta}}(X). \tag{5.5}$$

We may assume that O is given by

$$O = \Big\{ Q : \big| \int g_i(y)\, Q(dy) \big| < \epsilon_i,\ i = 1, \dots, K \Big\}$$

for bounded functions $g_i : \tilde{E}^{\mathbb{Z}} \to \mathbb{R}$ depending on finitely many coordinates and $\epsilon_i > 0$. By shift-invariance, we can and shall assume that $g_1, \dots, g_K$ depend only $y^{(0)}, \dots, y^{(M)}$ for some $M \in \mathbb{N}$, and we can equivalently treat g_i as a function on $\tilde{E}^{M+1}$.

For $x \in E^{\mathbb{Z}}$, M as above, $N \in \mathbb{N}$, $0 \le j_0 < j_1 < \dots < j_{N+M}$ write

$$\tilde{R}^{N,M}_{j_0,j_1,\dots,j_{N+M}}(x) := \frac{1}{N} \sum_{i=0}^{N-1} \delta_{\left(x|_{j_i,\dots,j_{i+1}-1},\, x|_{j_{i+1},\dots,j_{i+2}-1},\, \dots,\, x|_{j_{i+M-1},\dots,j_{i+M}-1}\right)}$$

for the element of $\mathcal{P}(\tilde{E}^{M+1})$ one obtains by cutting x at the given cut-points j_i and then building the empirical measure of order N of M-tuples out of the result. This allows to rewrite

$$\begin{aligned} \mathbb{P}(R_N \in O \,|\, X) &= \sum_{0 \le j_0 < j_1 < \dots < j_{N+M}} \prod_{i=0}^{N+M} \rho_{j_i - j_{i-1}} \mathbf{1}_{O^{(M)}}\big(\tilde{R}^{N,M}_{j_0,j_1,\dots,j_{N+M}}(X)\big) \\ &=: F_{N,O}(X), \end{aligned} \tag{5.6}$$

where

$$O^{(M)} = \Big\{ Q \in \mathcal{P}(\tilde{E}^{M+1}) : \big| \int g_i(y)\, Q(dy) \big| < \epsilon_i,\ i = 1, \dots, K \Big\}$$

is the projection of O to $\mathcal{P}(\tilde{E}^{M+1})$. Note that the random variable $\liminf \frac{1}{N} \log F_{N,O}(X)$ is adapted to the tail-σ-algebra of X, and thus almost surely constant. This implies

$$\liminf_{N\to\infty} \frac{1}{N} \log F_{N,O}(X) = \liminf_{N\to\infty} \frac{1}{N} \log F_{N,O}(\theta^{-1} X) \quad \text{a.s.} \tag{5.7}$$

(as there is a $c \in \overline{\mathbb{R}}$ and set $B \subset E^{\mathbb{Z}}$ of full $\nu^{\otimes Z}$-measure on which $\liminf \frac{1}{N} \log F_{N,O}(X) = c$, and $\nu^{\otimes Z} = \nu^{\otimes Z} \circ \theta$).

Applying (5.7) $|\tilde{\Delta}|$ times, in view of (5.5), we obtain

$$\begin{aligned} &\liminf \frac{1}{N} \log \mathbb{P}(R_N \in O \,|\, X) = \liminf \frac{1}{N} \log F_{N,O}(X) \\ &= \liminf \frac{1}{N} \log F_{N,O}(\theta^{\tilde{\Delta}} X) = \liminf \frac{1}{N} \log F_{N,O}\big(\kappa(Y)\big) \\ &\ge \liminf \frac{1}{N} \log \sum_{0 \le j_0 < j_1 < \dots < j_{N+M} \le (m_Q+\epsilon)(N+M)} \prod_{i=0}^{N+M} \rho_{j_i - j_{i-1}} \mathbf{1}_{O^{(M)}}\big(\tilde{R}^{N,M}_{j_0,j_1,\dots,j_{N+M}}(\kappa(Y))\big) \\ &\ge \liminf \frac{1}{N} \log \mathbb{E}_{\hat{Q}}\Big[\mathbf{1}_{\tilde{O}}(R_N) \mathbf{1}(L_0 + \dots + L_{N+M} \le (m_Q+\epsilon)(N+M)) \\ &\qquad \times \exp\Big(-\log \hat{Q}(L_0, \dots, L_{N+L-1} | \kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)}) \\ &\qquad\qquad + \sum_{i=0}^{N+M} \log \rho_{L_i} \Big) \Big| \kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)} \Big]. \end{aligned}$$

Furthermore:

$$\lim_{N\to\infty}\frac{1}{N}\log\frac{\hat{Q}(L_0,\dots,L_{N-1}|\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}{Q(L_0,\dots,L_{N-1}|\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}=0 \tag{5.8}$$

almost surely under Q (and hence also under $\hat{Q}$). To see this observe that the quotient inside the log is equal to

$$\frac{\hat{Q}(L_0,\dots,L_{N-1},\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}{Q(L_0,\dots,L_{N-1},\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}\times\frac{Q(\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}{\hat{Q}(\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}.$$

The first factor above is $L_0/\mathbb{E}_Q\,L_0$ for all N, the second is a fixed strictly positive random variable — $\mathscr{L}_Q(\kappa(Y))$ and $\mathscr{L}_{\hat{Q}}(\kappa(Y))$ are mutually absolutely continuous — so taking logarithms and dividing by N yields 0 as $N\to\infty$.)

By Lemma 7,

$$\limsup_{N\to\infty}\frac{1}{N}\log\frac{Q(L_0,\dots,L_{N-1}|\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}{\prod_{i=0}^{N+M-1}\rho_{L_i}}\le H(Q;Q^0)+\delta(\epsilon) \tag{5.9}$$

Q-almost surely (and hence also $\hat{Q}$-almost surely). Furthermore, $R_N\to Q$ (Q- and $\hat{Q}$-almost surely). Combining, we can bound

$$F_{N,\tilde{O}}(\kappa(Y))\quad\ge\quad\exp\big(-N(H(Q;Q^0)+2\epsilon+\delta(\epsilon))\big)\,\hat{Q}(A_{N,\epsilon}|\kappa(Y)),$$

where

$$A_{N,\epsilon}=\left\{\begin{array}{l}\left|\frac{1}{N}\log\frac{\hat{Q}(L_0,\dots,L_{N-1}|\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}{Q(L_0,\dots,L_{N-1}|\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}\right|\le\epsilon,\ L_0+\cdots+L_{N+M}\le(m_Q+\epsilon)(N+M),\\ \log\frac{Q(L_0,\dots,L_{N-1}|\kappa(Y)|_{0\dots(m_Q+\epsilon)(N+M)})}{\prod_{i=0}^{N+M-1}\rho_{L_i}}\le H(Q;Q^0)+\epsilon+\delta(\epsilon),\ R_N\in O\end{array}\right\}.$$

As $\hat{Q}(A_{N,\epsilon}|\kappa(Y))\to_{N\to\infty}1$ (because almost sure limits remain unchanged under typical conditioning), we obtain $\liminf\frac{1}{N}\log\mathbb{P}(R_N\in O\,|\,X)\ge -H(Q;Q^0)-2\epsilon-\delta(\epsilon)$. □

6 Proof of Theorem 1 and Corollary 1

Proof of Thm. 1. Taking $M\to\infty$ in Lemma 5 yields (1.7) for any fixed $F\subset\mathcal{P}(\tilde{E}^{\otimes\mathbb{Z}})$ closed. Using the fact that the weak topology on $\mathcal{P}(\tilde{E}^{\otimes\mathbb{Z}})$ is countably generated it is standard to strengthen the bound to hold with probability one simultaneously for all closed F, see e.g. [2], proof of Prop. III.2, (2).

For fixed open $G\subset\mathcal{P}(\tilde{E}^{\otimes\mathbb{Z}})$, (1.8) follows from Lemma 8: note that we can restrict to $Q\in G\cap\mathscr{R}\cap\mathcal{P}^{\mathrm{erg}}(\tilde{E}^{\mathbb{Z}})$ with $H(Q;Q^0)<\infty$, for otherwise the lower bound is trivial. For such Q, we have $\mathbb{E}_Q\,L_0<\infty$ by (3.3), so that Lemma 8 applies. Again, we can strengthen the statement to all open sets simultaneously as in [2]. □

Proof of Cor. 1. (1.9) is essentially Varadhan's Lemma applied to the scenario at hand. The only twist is that in the lower bound (1.8), we have the restriction that $Q\in\mathcal{P}^{\mathrm{erg}}(\tilde{E}^{\mathbb{Z}})$. Note that the rate function

$$Q\mapsto\begin{cases}H(Q;Q^0) & \text{if } Q\in\mathscr{R}\\ \infty & \text{otherwise}\end{cases}$$

has compact level sets by Lemma 3. Thus the upper half of (1.9), namely

$$\limsup_N \frac{1}{N} \log \mathbb{E}\Big[\exp\big(N \int \Phi(y^{(\cdot)}) R_N(dy^{(\cdot)})\big)\Big| X\Big] \le \sup_{Q \in \mathscr{R} \cap \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})} \{\dots\}$$

is standard, use e.g. Lemma 4.3.6 in [3] in conjunction with (1.7).

Using (1.8) in the standard proof of the lower half (e.g. Lemma 4.3.4 in [3]) we obtain

$$\liminf_N \frac{1}{N} \log \mathbb{E}\Big[\cdots \Big| X\Big] \ge \sup_{Q \in \mathscr{R} \cap \mathcal{P}^{\mathrm{erg}}(\tilde{E}^{\mathbb{Z}})} \Big\{ \int \Phi(y^{(\cdot)}) Q(dy^{(\cdot)}) - H(Q; Q^0) \Big\} \tag{6.1}$$

to begin with. The mapping $Q \mapsto \int \Phi \, dQ$ is linear, $Q \mapsto H(Q; Q^0)$ is affine (see e.g. [4], (5.4.23)), so using the decomposition of $Q \in \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})$ into ergodic components (e.g. [4], Thm. 5.2.16) we see that the right-hand side of (6.1) is in fact equal to

$$\sup_{Q \in \mathscr{R} \cap \mathcal{P}^{\mathrm{shift}}(\tilde{E}^{\mathbb{Z}})} \Big\{ \int \Phi(y^{(\cdot)}) Q(dy^{(\cdot)}) - H(Q; Q^0) \Big\},$$

completing the proof of (1.9). □

Acknowledgement

The author would like to thank Andreas Greven and Frank den Hollander for very stimulating discussions on this subject.

References

[1] Erwin Bolthausen, Frank den Hollander, Localization transition for a polymer near an interface. *Ann. Probab.* **25** (1997), no. 3, 1334–1366.

[2] Francis Comets, Large deviation estimates for a conditional probability distribution. Applications to random interaction Gibbs measures. *Probab. Theory Related Fields* **80** (1989), no. 3, 407–432.

[3] Amir Dembo, Ofer Zeitouni, *Large deviations techniques and applications.* Jones and Bartlett Publishers, Boston, 1993.

[4] Jean-Dominique Deuschel, Daniel W. Stroock, *Large deviations.* Academic Press, Boston, 1989.

[5] Claude Kipnis, Claudio Landim, *Scaling limits of interacting particle systems.* Springer-Verlag, Berlin, 1999.