A quenched large deviation principle in a continuous scenario

M. Birkner [1]
F. den Hollander [2]

8th December 2013

Abstract

We prove the analogue for continuous space-time of the quenched LDP derived in Birkner, Greven and den Hollander [2] for discrete space-time. In particular, we consider a random environment given by Brownian increments, cut into pieces according to an independent continuous-time renewal process. We look at the empirical process obtained by recording both the length of and the increments in the successive pieces. For the case where the renewal time distribution has a Lebesgue density with a polynomial tail, we derive the quenched LDP for the empirical process, i.e., the LDP conditional on a typical environment. The rate function is a sum of two specific relative entropies, one for the pieces and one for the concatenation of the pieces. We also obtain a quenched LDP when the tail decays faster than algebraic. The proof uses coarse-graining and truncation arguments, involving various approximations of specific relative entropies that are not quite standard.

In a companion paper we show how the quenched LDP and the techniques developed in the present paper can be applied to obtain a variational characterisation of the free energy and the phase transition line for the Brownian copolymer near a selective interface.

MSC2010: 60F10, 60G10, 60J65, 60K37.
Keywords: Brownian environment, renewal process, annealed vs. quenched, empirical process, large deviation principle, specific relative entropy.
Acknowledgment: The research in this paper is supported by ERC Advanced Grant 267356 VARIS of FdH. MB is grateful for hospitality at the Mathematical Institute in Leiden during a sabbatical leave from September 2012 until February 2013, supported by ERC.

1 Introduction and main result

When we cut an i.i.d. sequence of letters into words according to an independent integer-valued renewal process, we obtain an i.i.d. sequence of words. In the *annealed* LDP for the empirical process of words, the rate function is the specific relative entropy of the observed law of words w.r.t. the reference law of words. Birkner, Greven and den Hollander [2] considered the *quenched* LDP, i.e., conditional on a typical letter sequence. The rate function of the quenched LDP turned out to be a sum of two terms, one being the annealed rate function, the other being proportional to the specific relative entropy of the observed law of letters w.r.t. the reference law of letters, with the former being obtained by concatenating the words and randomising the location of the origin. The proportionality constant equals the tail exponent of the renewal time distribution.

[1] Institut für Mathematik, Johannes-Gutenberg-Universität, Staudingerweg 9, 55099 Mainz, Germany, *birkner@mathematik.uni-mainz.de*
[2] Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands, *denholla@math.leidenuniv.nl*

The goal of the present paper is to derive the analogue of the quenched LDP for the case where the i.i.d. sequence of letters is replaced by the process of Brownian increments, and the renewal process has a length distribution with a Lebesgue density that has a polynomial tail.

In Section 1.1 we define the continuous space-time setting, in Section 1.2 we state both the annealed and the quenched LDP, while in Section 1.3 we discuss these LDPs and indicate some further extensions. In Section 2 we prove the quenched LDP subject to three propositions. In Sections 3–4 we give the proof of these propositions. In Section 5 we prove the extensions. Appendix A recalls a few basic facts about metrics on path space, while Appendices B–C prove a few basic facts about specific relative entropy that are needed in the proof and that are not quite standard.

1.1 Continuous space-time

Let $X = (X_t)_{t\geq 0}$ be the standard one-dimensional Brownian motion starting from $X_0 = 0$. Let $\mathscr{W}$ denote its law on path space: the Wiener measure on $C([0,\infty))$, equipped with the σ-algebra generated by the coordinate projections. Let $T = (T_i)_{i\in\mathbb{N}_0}$ ($T_0 = 0$) be an independent continuous-time renewal process, with interarrival times $\tau_i = T_i - T_{i-1}$, $i \in \mathbb{N}$, whose common law $\rho = \mathscr{L}(\tau_1)$ is absolutely continuous with respect to the Lebesgue measure on $(0,\infty)$, with density $\bar\rho$ satisfying

$$\lim_{x\to\infty} \frac{\log \bar\rho(x)}{\log x} = -\alpha, \qquad \alpha \in (1,\infty). \tag{1.1}$$

In addition, assume that

$$\text{supp}(\rho) = [s_*,\infty) \text{ with } 0 \leq s_* < \infty, \text{ and } \bar\rho \text{ is continuous and strictly positive on } (s_*,\infty), \text{ and varies regularly near } s_*. \tag{1.2}$$

Define the *word sequence* $Y = (Y^{(i)})_{i\in\mathbb{N}}$ by putting (see Fig. 1)

$$Y^{(i)} = \Big(T_i - T_{i-1}, \big(X_{(s+T_{i-1})\wedge T_i} - X_{T_{i-1}}\big)_{s\geq 0}\Big), \tag{1.3}$$

which takes values in the *word space*

$$F = \bigcup_{t>0} \Big(\{t\} \times \big\{f \in C([0,\infty))\colon\ f(0) = 0, f(s) = f(t) \text{ for } s > t\big\}\Big) \tag{1.4}$$

equipped with a Skorohod-type metric (see Appendix A). Let

$$Y^{N\text{-per}} = \big(\, \underbrace{Y^{(1)}, Y^{(2)}, \dots, Y^{(N)}}, \underbrace{Y^{(1)}, Y^{(2)}, \dots, Y^{(N)}}, \dots\big) \tag{1.5}$$

denote the N-periodisation of Y, and let

$$R_N = \frac{1}{N} \sum_{i=0}^{N-1} \delta_{\widetilde\theta^i Y^{N\text{-per}}} \tag{1.6}$$

be the *empirical process of words*, where $\widetilde\theta$ is the left-shift acting on $F^{\mathbb{N}}$. Note that R_N takes values in $\mathcal{P}^{\text{inv}}(F^{\mathbb{N}})$, the set of shift-invariant probability measures on $F^{\mathbb{N}}$. Endow $F^{\mathbb{N}}$ with the product topology and $\mathcal{P}^{\text{inv}}(F^{\mathbb{N}})$ with the corresponding weak topology. When averaged over X and T, the law of Y is ($\mathscr{L}$ denotes law)

$$Q_{\rho,\mathscr{W}} = (q_{\rho,\mathscr{W}})^{\otimes\mathbb{N}} \quad \text{with} \quad q_{\rho,\mathscr{W}} = \int_{(0,\infty)} \rho(dt)\, \mathscr{L}\big((t, (X_{s\wedge t})_{s\geq 0})\big). \tag{1.7}$$

By the ergodic theorem, $\text{w-lim}_{N\to\infty} R_N = Q_{\rho,\mathscr{W}}$ a.s., where w-lim denotes the weak limit.

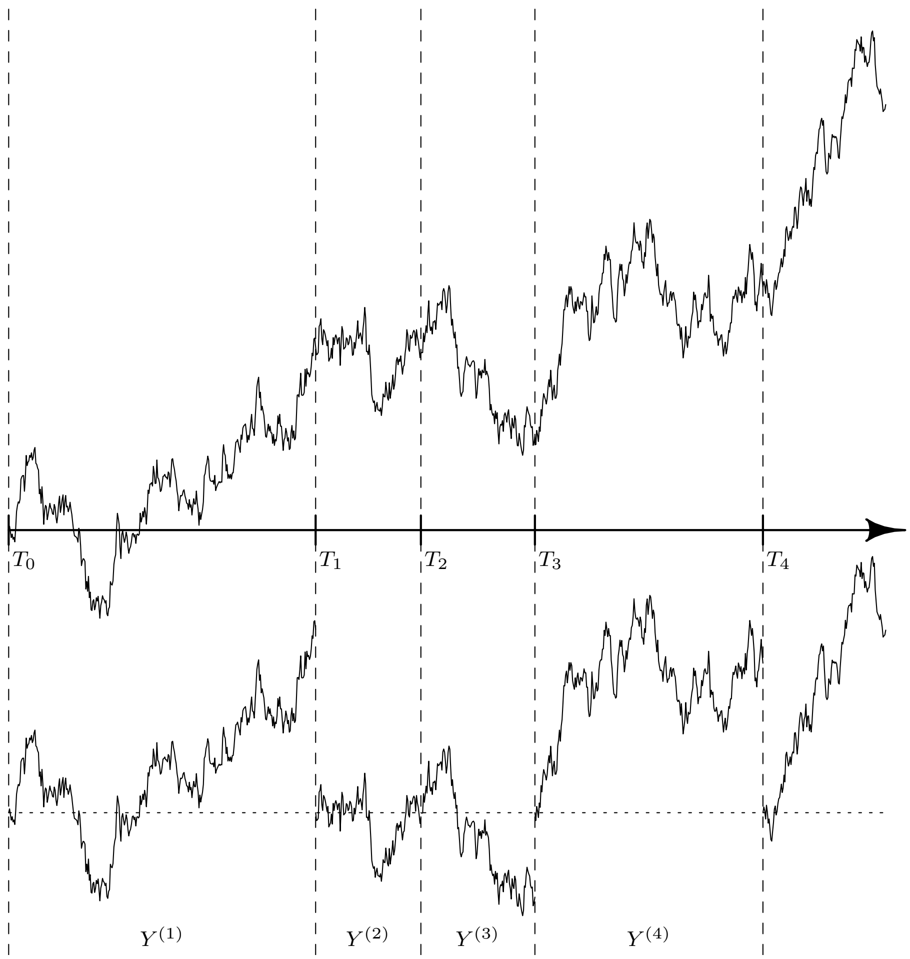

Figure 1: The word sequence Y. Upper part: Brownian path X and renewal times T. Lower part: increments of the path between the renewal times (which are elements of F).

1.2 Large deviation principles

For definitions and properties of specific relative entropy, we refer the reader to Appendix B.

The following theorem is standard (see e.g. Dembo and Zeitouni [7, Section 6.5.3]).

Theorem 1.1. [Annealed LDP]
The family $\mathscr{L}(R_N)$, $N \in \mathbb{N}$, satisfies the LDP on $\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with rate N and with rate function

$$I^{\mathrm{ann}}(Q) = H(Q \mid Q_{\rho,\mathscr{W}}), \tag{1.8}$$

the specific relative entropy of Q w.r.t. $Q_{\rho,\mathscr{W}}$. This rate function is lower semi-continuous, has compact level sets, is affine, and has a unique zero at $Q = Q_{\rho,\mathscr{W}}$.

To state the quenched LDP, we need to look at the reverse of cutting out words, namely, glueing words together. Let $y = (y^{(i)})_{i\in\mathbb{N}} = ((t_i, f_i))_{i\in\mathbb{N}} \in F^{\mathbb{N}}$. Then the *concatenation* of y, written $\kappa(y) \in C([0,\infty))$, is defined by

$$\begin{aligned} &\kappa(y)(s) = f_1(t_1) + \cdots + f_{i-1}(t_{i-1}) + f_i\big(s - (t_1 + \cdots + t_{i-1})\big), \\ &t_1 + \cdots + t_{i-1} \le s < t_1 + \cdots + t_i, \ \ i \in \mathbb{N}. \end{aligned} \tag{1.9}$$

Write $\tau_i(y) = t_i$ to denote the length of the i-th word. For $Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with finite mean word length $m_Q = \mathbb{E}_Q[\tau_1] = \mathbb{E}_Q[\tau_1(Y)]$, put

$$\Psi_Q(A) = \frac{1}{m_Q}\mathbb{E}_Q\left[\int_0^{\tau_1} 1_A(\theta^s \kappa(Y))\, ds\right], \quad A \subset C([0,\infty)) \text{ measurable}, \tag{1.10}$$

where θ^s is the shift acting on $f \in C([0,\infty))$ as $\theta^s f(t) = f(s+t) - f(s)$, $t \ge 0$. Note that Ψ_Q is a probability measure on $C([0,\infty))$ with stationary increments, i.e., $\Psi_Q = \Psi_Q \circ (\theta^s)^{-1}$ for all $s \ge 0$. We can think of Ψ_Q as the "stationarised" version of $\kappa(Q)$. In fact, if $m_Q < \infty$, then

$$\Psi_Q = \operatorname*{w-lim}_{T\to\infty} \frac{1}{T}\int_0^T \kappa(Q) \circ (\theta^s)^{-1}\, ds, \tag{1.11}$$

and $\kappa(Q)$ is asymptotically mean stationary (AMS) with stationary mean Ψ_Q. In fact, the convergence in (1.11) also holds in total variation norm (see Lemma B.4 in Appendix B). Note that $\Psi_{Q_{\rho,\mathscr{W}}} = \mathscr{W}$.

To state the quenched LDP, we also need to define word *truncation*. For $(t,f) \in F$ and $\mathrm{tr} > 0$, let

$$[(t,f)]_{\mathrm{tr}} = \big(t \wedge \mathrm{tr}, (f(s \wedge \mathrm{tr})_{s\ge 0}\big) \tag{1.12}$$

be the word (t,f) truncated at length tr. Analogously, for $y = (y^{(i)})_{i\in\mathbb{N}} \in F^{\mathbb{N}}$ set $[y]_{\mathrm{tr}} = ([y^{(i)}]_{\mathrm{tr}})_{i\in\mathbb{N}} \in F^{\mathbb{N}}$, and denote by $[Q]_{\mathrm{tr}} \in \mathcal{P}^{\mathrm{inv}}(F_{0,\mathrm{tr}}^{\mathbb{N}}) \subset \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with $F_{0,\mathrm{tr}} = [F]_{\mathrm{tr}}$ the image measure of $Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ under the map $y \mapsto [y]_{\mathrm{tr}}$.

Theorem 1.2. [Quenched LDP]
Suppose that ρ satisfies (1.1–1.2). Then, for $\mathscr{W}$ a.e. X, the family $\mathscr{L}(R_N \mid X)$, $N \in \mathbb{N}$, satisfies the LDP on $\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with rate N and with deterministic rate function $I^{\mathrm{que}}(Q)$ given by

$$I^{\mathrm{que}}(Q) = \lim_{\mathrm{tr}\to\infty} I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}}), \tag{1.13}$$

where

$$I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}}) = H\big([Q]_{\mathrm{tr}} \mid [Q_{\rho,\mathscr{W}}]_{\mathrm{tr}}\big) + (\alpha - 1) m_{[Q]_{\mathrm{tr}}} H\big(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}\big). \tag{1.14}$$

This rate function is lower semi-continuous, has compact level sets, is affine, and has a unique zero at $Q = Q_{\rho,\mathscr{W}}$.

Theorem 1.2 is proved in Sections 2–4. Let $\mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}}) = \{Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})\colon\, m_Q < \infty\}$. We will show that the limit in (1.13) exists for all $Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$, and that

$$I^{\mathrm{que}}(Q) = H(Q \mid Q_{\rho,\mathscr{W}}) + (\alpha-1) m_Q H(\Psi_Q \mid \mathscr{W}), \qquad Q \in \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}}). \tag{1.15}$$

We will also see that $I^{\mathrm{que}}(Q)$ is the lower semi-continuous extension to $\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ of its restriction to $\mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$.

1.3 Discussion

0. A *heuristic* behind Theorem 1.2 is as follows. Let

$$R^N_{t_1,\dots,t_N}(X), \qquad 0 < t_1 < \cdots < t_N < \infty, \tag{1.16}$$

denote the empirical process of N-tuples of words when X is cut at the points $t_1,\dots,t_N$ (i.e., when $T_i = t_i$ for $i = 1,\dots,N$). Fix $Q \in \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$ and suppose that Q is shift-ergodic. The probability $\mathbb{P}(R_N \approx Q \mid X)$ is an integral over all N-tuples $t_1,\dots,t_N$ such that $R^N_{t_1,\dots,t_N}(X) \approx Q$, weighted by $\prod_{i=1}^N \bar\rho(t_i - t_{i-1})$ (with $t_0 = 0$). The fact that $R^N_{t_1,\dots,t_N}(X) \approx Q$ has three consequences:

(1) The $t_1,\dots,t_N$ must cut $\approx N$ substrings out of X of total length $\approx N m_Q$ that look like the concatenation of words that are Q-typical, i.e., that look as if generated by Ψ_Q (possibly with gaps in between). This means that most of the cut-points must hit atypical pieces of X. We expect to have to shift X by $\approx \exp[N m_Q H(\Psi_Q \mid \mathscr{W})]$ in order to find the first contiguous substring of length $N m_Q$ whose empirical shifts lie in a small neighbourhood of Ψ_Q. By (1.1), the probability for the single increment $t_1 - t_0$ to have the size of this shift is $\approx \exp[-N\alpha\, m_Q H(\Psi_Q \mid \mathscr{W})]$.

(2) The "number of local perturbations" of $t_1,\dots,t_N$ preserving the property $R^N_{t_1,\dots,t_N}(X) \approx Q$ is $\approx \exp[N H_{\tau|K}(Q)]$, where $H_{\tau|K}$ stands for the *conditional specific entropy (density) of word lengths under the law Q*.

(3) The statistics of the increments $t_1 - t_0,\dots,t_N - t_{N-1}$ must be close to the distribution of word lengths under Q. Hence, the weight factor $\prod_{i=1}^N \bar\rho(t_i - t_{i-1})$ must be $\approx \exp[N \mathbb{E}_Q[\log \bar\rho(\tau_1)]]$ (at least, for Q-typical pieces).

Since

$$m_Q H(\Psi_Q \mid \mathscr{W}) - H_{\tau|K}(Q) - \mathbb{E}_Q[\log \bar\rho(\tau_1)] = H(Q \mid q_{\rho,\mathscr{W}}), \tag{1.17}$$

the observations made in (1)–(3) combine to explain the shape of the quenched rate function in (1.15). For further details, see [2, Section 1.5].

Note: We have not defined $H_{\tau|K}(Q)$ rigorously here, nor do we prove (1.17). Our proof of Theorem 1.2 uses the above heuristic only very implicitly. Rather, it starts from the discrete-time quenched LDP derived in [2] and draws out Theorem 1.2 via control of exponential functionals through a coarse-graining approximation.

1. We can include the cases $\alpha = 1$ and $\alpha = \infty$ in (1.1).

Theorem 1.3. *Suppose that ρ satisfies* (1.1–1.2).
(a) *If $\alpha = 1$, then the quenched LDP holds with $I^{\mathrm{que}} = I^{\mathrm{ann}}$ given by* (1.8).
(b) *If $\alpha = \infty$, then the quenched LDP holds with rate function*

$$I^{\mathrm{que}}(Q) = \begin{cases} H(Q \mid Q_{\rho,\mathscr{W}}) & \text{if } \lim_{\mathrm{tr}\to\infty} m_{[Q]_{\mathrm{tr}}} H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}) = 0, \\ \infty & \text{otherwise.} \end{cases} \tag{1.18}$$

Theorem 1.3 is the continuous analogue of Birkner, Greven and den Hollander [2, Theorem 1.4] and is proved in Section 5.

2. We can also include the case where $\bar{\rho}$ has an exponentially bounded tail:

$$\bar{\rho}(t) \le e^{-\lambda t} \text{ for some } \lambda > 0 \text{ and } t \text{ large enough.} \tag{1.19}$$

Theorem 1.4. *Suppose that ρ satisfies* (1.1–1.2) *and* (1.19)*. Then, for $\mathscr{W}$ a.e. X, the family $\mathscr{L}(R_N \mid X)$, $N \in \mathbb{N}$, satisfies the LDP on $\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with rate N and with deterministic rate function $I^{\mathrm{que}}(Q)$ given by*

$$I^{\mathrm{que}}(Q) = \begin{cases} H(Q \mid Q_{\rho,\mathscr{W}}) & \text{if } Q \in \mathcal{R}_{\mathscr{W}}, \\ \infty & \text{otherwise,} \end{cases} \tag{1.20}$$

where

$$\mathcal{R}_{\mathscr{W}} = \left\{ Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})\colon \ \operatorname*{w-lim}_{T\to\infty} \frac{1}{T} \int_0^T \delta_{\kappa(Y)} \circ (\theta^s)^{-1}\, ds = \mathscr{W} \ \text{ for } Q\text{-a.e. } Y \right\}. \tag{1.21}$$

Theorem 1.4 is the continuous analogue of Birkner [1, Theorem 1] and is proved in Section 5. On the set $\mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$ the following holds:

$$\Psi_Q = \mathscr{W} \quad \text{if and only if} \quad Q \in \mathcal{R}_{\mathscr{W}}. \tag{1.22}$$

The equivalence in (1.22) is the continuous analogue of [1, Lemma 2] (and can be proved analogously).

3. By applying the contraction principle we obtain the quenched LDP for single words. Let $\pi_1\colon F^{\mathbb{N}} \to F$ be the projection onto the first word, and let $\pi_1 R_N = R_N \circ (\pi_1)^{-1}$.

Corollary 1.5. *Suppose that ρ satisfies* (1.1–1.2)*. For $\mathscr{W}$-a.e. X, the family $\mathscr{L}(\pi_1 R_N \mid X)$, $N \in \mathbb{N}$, satisfies the LDP on $\mathcal{P}(F)$ with rate N and with deterministic rate function I_1^{que} given by*

$$I_1^{\mathrm{que}}(q) = \inf\left\{ I^{\mathrm{que}}(Q)\colon\ Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}}),\ \pi_1 Q = q \right\}. \tag{1.23}$$

This rate function is lower semi-continuous, has compact levels sets, is convex, and has a unique zero at $q = q_{\rho,\mathscr{W}}$.

For general q it is not possible to evaluate the infimum in (1.23) explicitly. For q with $m_q = \mathbb{E}_q[\tau] = \mathbb{E}_{q^{\otimes\mathbb{N}}}[\tau_1] = m_{q^{\otimes\mathbb{N}}} < \infty$ and $\Psi_{q^{\otimes\mathbb{N}}} = \mathscr{W}$, we have $I_1^{\mathrm{que}}(q) = h(q \mid q_{\rho,\mathscr{W}})$, the relative entropy of q w.r.t. $q_{\rho,\mathscr{W}}$.

4. We expect assumption (1.2) to be redundant. In any case, it can be relaxed to (see Section 3.1):

$\operatorname{supp}(\rho) = \cup_{i=1}^M [a_i, b_i] \cup [a_{M+1}, \infty)$ with $M \in \mathbb{N}$ and $0 \le a_1 < b_1 \le a_2 < \cdots < b_M \le a_{M+1} < \infty$, and $\bar{\rho}$ is continuous and strictly positive on $\cup_{i=1}^M (a_i, b_i) \cup (a_{M+1}, \infty)$ and varies regularly near each of the finite endpoints of these intervals. (1.24)

5. It is possible to extend Theorem 1.2 to other classes of random environments, as stated in the following theorem whose proof will not be spelled out in the present paper.

Theorem 1.6. *Theorems* 1.2–1.4 *and Corollary* 1.5 *carry over verbatim when the Brownian motion X is replaced by a d-dimensional Lévy process $\bar{X}$ with the property that $\mathbb{E}[e^{\langle \lambda, \bar{X}_1 \rangle}] < \infty$ for all $\lambda \in \mathbb{R}^d$ (where $\langle \cdot \rangle$ denotes the standard inner product), $\mathscr{W}$ is replaced by the law of $\bar{X}$, and in the definition of F in* (1.4) *continuous paths are replaced by càdlàg paths.*

6. In the companion paper [3] we apply Theorem 1.2 and the techniques developed in the present paper to the Brownian copolymer. In this model a càdlàg path, representing the configuration of the polymer, is rewarded or penalised for staying above or below a linear interface, separating oil from water, according to Brownian increments representing the degrees of hydrophobicity or hydrophilicity along the polymer. The reference measure for the path can be either the Wiener measure or the law of a more general Lévy process. We derive a variational formula for the quenched free energy, from which we deduce a variational formula for the slope of the quenched critical line. This critical line separates a *localized phase* (where the copolymer stays close to the interface) from a *delocalized phase* (where the copolymer wanders away from the interface). This slope has been the object of much debate in recent years. The Brownian copolymer is the unique attractor in the limit of weak interaction for a whole universality class of discrete copolymer models. See Bolthausen and den Hollander [4], Caravenna and Giacomin [5], Caravenna, Giacomin and Toninelli [6] for details.

2 Proof of Theorem 1.2

The proof proceeds via a *coarse-graining* and *truncation* argument. In Section 2.1 we set up the coarse-graining and the truncation, and state a quenched LDP for this setting that follows from the quenched LDP in [2] and serves as the starting point of our analysis (Proposition 2.1 and Corollary 2.2 below). In Section 2.2 we state three propositions (Propositions 2.3–2.5 below), involving expectations of exponential functionals of the coarse-grained truncated empirical process as well as approximation properties of the associated rate function, and we use these propositions to complete the proof of Theorem 1.2 with the help of Bryc's inverse of Varadhan's lemma. In Section 2.3 we state and prove two lemmas that are used in Section 2.2, involving approximation estimates under the coarse-graining. The proof of the three propositions is deferred to Sections 3–4.

2.1 Preparation: coarse-graining and truncation

2.1.1 Coarse-graining

Suppose that, instead of the absolutely continuous ρ introduced in Section 1.1, we are given a discrete $\hat{\rho}$ with $\mathrm{supp}(\hat{\rho}) \subset h\mathbb{N}$ for some $h > 0$. Let

$$E_h = \{f \in C([0,h])\colon\ f(0) = 0\}. \tag{2.1}$$

Path pieces of length h in a continuous-time scenario can act as "letters" in a discrete-time scenario, and therefore we can use the results from [2]. Note that $(E_h)^{\mathbb{N}}$ as a metric space is isomorphic to $\{f \in C([0,\infty))\colon\ f(0) = 0\}$ via the obvious glueing together of path pieces into a single path, provided the latter is given a suitable metric that metrises locally uniform convergence. Similarly, we can identify $\mathcal{P}^{\mathrm{inv}}(E_h^{\mathbb{N}})$ with

$$\mathcal{P}^{h\text{-inv}}(C([0,\infty))) = \big\{Q \in \mathcal{P}(C([0,\infty)))\colon\ Q = Q \circ (\theta^h)^{-1}\big\}, \tag{2.2}$$

which is the set of laws on continuous paths that are invariant under a time shift by h. Note that the set

$$F_h = \bigcup_{t \in h\mathbb{N}} \Big(\{t\} \times \big\{ f \in C([0,\infty))\colon\ f(0) = 0, f(s) = f(t) \text{ for } s > t \big\} \Big) \tag{2.3}$$

is isomorphic to $\widetilde{E_h} = \cup_{n\in\mathbb{N}} (E_h)^n$ via the map $\iota_h\colon F_h \to \widetilde{E_h}$ defined by

$$\iota_h\big((nh, f)\big) = \Big(\big(f\big((\cdot + (i-1)h) \wedge ih \big) - f((i-1)h) \big) \Big)_{i=1,\dots,n}, \qquad (nh, f) \in F_h. \tag{2.4}$$

For $Q \in \mathcal{P}^{\mathrm{inv,fin}}(F_h^{\mathbb{N}})$, define

$$\Psi_{Q,h}(A) = \frac{1}{m_Q} \mathbb{E}_Q \left[\sum_{i=0}^{\tau_1 - 1} 1_A\big(\theta^i \iota_h \kappa(Y) \big) \right] = \frac{1}{h\, m_Q} \mathbb{E}_Q \left[\int_0^{h\tau_1} 1_A\big(\kappa(Y)(h\lfloor u/h \rfloor + s))_{s \geq 0} \big)\, du \right] \tag{2.5}$$

for $A \subset C([0,\infty))$ measurable, where τ_1 is the length of the first word (counted in letters, so that the length of the first word viewed as an element of F_h is $h\tau_1$) and θ is the left-shift acting on $(E_h)^{\mathbb{N}}$. The right-most expression in (2.5) can be viewed as a coarse-grained version of (1.10). The following coarse-grained version of the quenched LDP serves as our starting point.

Proposition 2.1. *Fix $h > 0$. Suppose that* $\mathrm{supp}(\hat\rho) \subset h\mathbb{N}$ *and* $\lim_{n\to\infty} \log \hat\rho(\{nh\}) / \log n = -\alpha$ *with $\alpha \in (1,\infty)$. Then, for $\mathscr{W}$ a.e. X, the family $\mathscr{L}(R_N \mid X)$, $N \in \mathbb{N}$, satisfies the LDP on $\mathcal{P}^{\mathrm{inv}}((\widetilde{E_h})^{\mathbb{N}})$ with rate N and with deterministic rate function given by*

$$I_h^{\mathrm{que}}(Q) = H(Q \mid Q_{\hat\rho,\mathscr{W}}) + (\alpha - 1) m_Q H(\Psi_{Q,h} \mid \mathscr{W}), \qquad Q \in \mathcal{P}^{\mathrm{inv,fin}}((\widetilde{E_h})^{\mathbb{N}}), \tag{2.6}$$

and

$$I_h^{\mathrm{que}}(Q) = \lim_{\mathrm{tr}\to\infty} I_h^{\mathrm{que}}([Q]_{\mathrm{tr}}), \qquad Q \notin \mathcal{P}^{\mathrm{inv,fin}}((\widetilde{E_h})^{\mathbb{N}}), \tag{2.7}$$

where $Q_{\hat\rho,\mathscr{W}} = (q_{\hat\rho,\mathscr{W}})^{\otimes\mathbb{N}}$ with $q_{\hat\rho,\mathscr{W}}$ defined as in (1.7), and $\Psi_{Q,h}$ defined via (2.5).

Proof. The claim follows from [2, Corollary 1.6] by using E_h as letter space and observing that $\widetilde{E_h} = \iota_h(F_h)$. Note that $F_h^{\mathbb{N}}$ is a closed subspace of $F^{\mathbb{N}}$. Since $\mathrm{supp}(\hat\rho) \subset h\mathbb{N}$ by assumption, we have $I_h^{\mathrm{que}}(Q) \geq H(Q \mid Q_{\hat\rho,\mathscr{W}}) = \infty$ for any $Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with $Q\big(F^{\mathbb{N}} \backslash F_h^{\mathbb{N}}\big) > 0$. Therefore we can consider the random variable R_N as taking values in $\mathcal{P}^{\mathrm{inv}}((\widetilde{E_h})^{\mathbb{N}})$, $\mathcal{P}^{\mathrm{inv}}(F_h^{\mathbb{N}})$ or $\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$, without changing the statement of Proposition 2.1. Note that I_h^{que} is finite only on $\mathcal{P}^{\mathrm{inv}}(F_h^{\mathbb{N}}) \subset \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$. □

We want to pass to the limit $h \downarrow 0$ and deduce Theorem 1.2 from Proposition 2.1. However, an immediate application of a projective limit at the level of letters appears to be impossible. Indeed, when we replace h by $h/2$, each "h-letter" turns into two "$(h/2)$-letters", so the word length changes, and even diverges as $h \downarrow 0$. This does not fit well with the way the projective limit was set up in [2, Section 8], where the internal structure of the letters was allowed to become increasingly richer, but the word length had to remain the same. In some sense, the problem is that we have finite words but only infinitesimal letters (i.e., there is no fixed letter space). To remedy this, we proceed as follows. For fixed discretisation length $h > 0$ we have a fixed letter space, and so Proposition 2.1 applies. We will handle the limit $h \downarrow 0$ via Bryc's inverse of Varadhan's lemma. This will require several intermediate steps.

2.1.2 Truncation

It will be expedient to work with a *truncated* version of Proposition 2.1. For $h > 0$, let $\lceil t \rceil_h = h\lceil t/h \rceil$ for $t \in (0,\infty)$ and put $\lceil \rho \rceil_h = \rho \circ (\lceil \cdot \rceil_h)^{-1}$, i.e.,

$$\lceil \rho \rceil_h = \sum_{i\in\mathbb{N}} w_{h,i}\delta_{ih} \in \mathcal{P}(h\mathbb{N}) \subset \mathcal{P}((0,\infty)), \tag{2.8}$$

where

$$w_{h,i} = \rho\big(((i-1)h, ih]\big) = \int_{(i-1)h}^{ih} \bar{\rho}(x)dx \tag{2.9}$$

is the coarse-grained version of ρ from Section 1.1. It is easily checked that (1.1) implies

$$\lim_{n\to\infty} \frac{\log \lceil \rho \rceil_h(\{nh\})}{\log n} = -\alpha. \tag{2.10}$$

Write$\mathscr{L}_{\lceil \rho \rceil_h}([R_N]_{\mathrm{tr}} \mid X)$ for the law of the truncated empirical process $[R_N]_{\mathrm{tr}}$ conditional on X when the τ_i's are drawn according to $\lceil \rho \rceil_h$.

Corollary 2.2. *For $\mathscr{W}$-a.e. X, the family $\mathscr{L}_{\lceil \rho \rceil_h}([R_N]_{\mathrm{tr}} \mid X)$, $N \in \mathbb{N}$, satisfies the LDP on $\mathcal{P}^{\mathrm{inv}}(F_h^{\mathbb{N}})$ with rate N and with deterministic rate function given by*

$$I^{\mathrm{que}}_{h,\mathrm{tr}}(Q) = H\big(Q \mid Q_{\lceil \rho \rceil_h,\mathscr{W},\mathrm{tr}}\big) + (\alpha-1)m_Q H(\Psi_{Q,h} \mid \mathscr{W}) \tag{2.11}$$

with $Q_{\lceil \rho \rceil_h,\mathscr{W},\mathrm{tr}} = ([q_{\lceil \rho \rceil_h,\mathscr{W}}]_{\mathrm{tr}})^{\otimes\mathbb{N}}$.

Proof. This follows from Proposition 2.1 and the contraction principle. Alternatively, it follows from the proofs of [2, Theorem 1.2 and Corollary 1.6]. □

Note that $I^{\mathrm{que}}_{h,\mathrm{tr}}(Q) = \infty$ when under Q the word lengths are not supported on $h\mathbb{N} \cap (0,\mathrm{tr}]$.

2.2 Application of Bryc's inverse of Varadhan's lemma

In this section we state three propositions (Propositions 2.3–2.5 below) and show that these imply Theorem 1.2. The proof of these propositions is deferred to Sections 3–4.

2.2.1 Notations

In what follows we obtain the quenched LDP for the truncated empirical process $[R_N]_{\mathrm{tr}}$ by letting $h \downarrow 0$ in the coarse-grained and truncated empirical process $[R_{N,h}]_{\mathrm{tr}}$ with $\mathrm{tr} \in \mathbb{N}$ fixed (for a precise definition, see (3.1) in Section 3.1) and afterwards letting $\mathrm{tr} \to \infty$. (We assume that $\mathrm{tr} \in \mathbb{N}$ and $h = 2^{-M}$ for some $M \in \mathbb{N}$, in particular, tr is an integer multiple of h.)

In the coarse-graining procedure, it may happen that a very short continuous word $y = (t,f) \in F$ disappears, namely, when $0 < t < h$. We remedy this by formally allowing "empty" words, i.e., by using

$$\widehat{F} = F \cup \big\{(0,0)\big\} = \bigcup_{t\geq 0} \Big(\{t\} \times \big\{f \in C([0,\infty))\colon\ f(0)=0, f(s)=f(t) \text{ for } s>t\big\}\Big) \tag{2.12}$$

as word space instead of F. The metric on F defined in Appendix A extends in the obvious way to $\widehat{F}$.

Before we proceed, we impose *additional regularity assumptions* on $\bar\rho$ that will be required in the proof of Proposition 2.3. Recall from (1.2) that $\mathrm{supp}(\rho) = [s_*, \infty)$. Let

$$V_{\bar\rho}(t,h) = \sup_{v\in(0,2h)} \left| \log \frac{\int_t^{t+h} \bar\rho(s)\,ds}{\int_{t+v}^{t+h+v} \bar\rho(s)\,ds} \right|, \qquad t,h>0. \tag{2.13}$$

We assume that there exist monotone sequences $(\eta_n)_{n\in\mathbb{N}}$ and $(A_n)_{n\in\mathbb{N}}$, with $\eta_n \in (0,1)$ and $A_n \subset (s_*,\infty)$ satisfying $\lim_{n\to\infty}\eta_n = 0$ and $\lim_{n\to\infty} A_n = (s_*,\infty)$, such that $(s_*,\infty)\setminus A_n$ is a (possibly empty) union of finitely many bounded intervals whose endpoints lie in $2^{-n}\mathbb{N}_0$, and

$$\sup_{t\in A_n} V_{\bar\rho}(t,2^{-n}) \le \eta_n \qquad \forall\, n\in\mathbb{N}. \tag{2.14}$$

In addition, we assume that there exists an $\eta_0 < \infty$ such that

$$\sup_{n\in\mathbb{N}} \sup_{t\in(s_*,\infty)} V_{\bar\rho}(t,2^{-n}) \le \eta_0. \tag{2.15}$$

These assumptions will be removed only in Section 4. Note that (2.14)–(2.15) are satisfied when $\bar\rho$ is continuous and strictly positive on (s_*,∞) and varies regularly near s_* and at ∞.

2.2.2 Proof of Theorem 1.2 subject to (2.14–2.15) and three propositions

Proof. A function g on $\widehat{F}^\ell$ is Lipschitz when it satisfies

$$\left|g(y^{(1)},\dots,y^{(\ell)}) - g(y^{(1)\prime},\dots,y^{(\ell)\prime})\right| \le C_g \sum_{j=1}^{\ell} d_F(y^{(j)},y^{(j)\prime}) \quad \text{for some } C_g<\infty. \tag{2.16}$$

Consider the class $\mathscr{C}$ of functions $\Phi\colon\ \mathcal{P}(\widehat{F}^{\mathbb{N}}) \to \mathbb{R}$ of the form

$$\Phi(Q) = \int_{\widehat{F}^{\ell_1}} g_1\, d\pi_{\ell_1}Q \wedge\cdots\wedge \int_{\widehat{F}^{\ell_m}} g_m\, d\pi_{\ell_m}Q, \quad Q\in\mathcal{P}^{\mathrm{inv}}(\widehat{F}^{\mathbb{N}}), \tag{2.17}$$

where $m\in N$, $\ell_1,\dots,\ell_m\in\mathbb{N}$, and g_i is a bounded Lipschitz function on $\widehat{F}^{\ell_i}$ for $i=1,\dots,m$. This class is well-separating and thus is sufficient for the application of Bryc's lemma (see Dembo and Zeitouni [7, Section 4.4].

Our first proposition identifies the exponential moments of $[R_N]_{\mathrm{tr}}$.

Proposition 2.3. *The families $\mathscr{L}(R_N \mid X)$, $N\in\mathbb{N}$, and $\mathscr{L}([R_N]_{\mathrm{tr}} \mid X)$, $\mathrm{tr}\in\mathbb{N}$, are exponentially tight X-a.s. Moreover, for $\Phi\in\mathscr{C}$,*

$$\Lambda_{0,\mathrm{tr}}(\Phi) = \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\Big[\exp\big(N\Phi([R_N]_{\mathrm{tr}})\big)\ \Big|\ X\Big] = \lim_{h\downarrow 0}\Lambda_{h,\mathrm{tr}}(\Phi) \quad \textit{exists } X\textit{-a.s.}, \tag{2.18}$$

where $\Lambda_{h,\mathrm{tr}}$ is the generalised convex transform of $I^{\mathrm{que}}_{h,\mathrm{tr}}$ given by

$$\Lambda_{h,\mathrm{tr}}(\Phi) = \sup_{Q\in\mathcal{P}^{\mathrm{inv,fin}}((\widetilde{E_h})^{\mathbb{N}})} \big\{\Phi(Q) - I^{\mathrm{que}}_{h,\mathrm{tr}}(Q)\big\}. \tag{2.19}$$

Furthermore, for $\Phi\in\mathscr{C}$,

$$\Lambda(\Phi) = \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}\Big[\exp\big(N\Phi(R_N)\big)\ \Big|\ X\Big] = \lim_{\mathrm{tr}\to\infty}\Lambda_{0,\mathrm{tr}}(\Phi) \quad \textit{exists } X\textit{-a.s.} \tag{2.20}$$

Our second proposition identifies the limit in (2.18) as the generalised convex transform of $I^{\mathrm{que}}_{\mathrm{tr}}$ defined in (1.14),

$$I^{\mathrm{que}}_{\mathrm{tr}}(Q) = \begin{cases} H\big(Q \mid Q_{\rho,\mathscr{W},\mathrm{tr}}\big) + (\alpha-1) m_Q H\left(\Psi_Q \mid \mathscr{W}\right) & \text{if } Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}}_{0,\mathrm{tr}}), \\ \infty & \text{otherwise}, \end{cases} \tag{2.21}$$

and implies that the latter is the rate function for the truncated empirical process $[R_N]_{\mathrm{tr}}$.

Proposition 2.4. *For $\Phi \in \mathscr{C}$,*

$$\Lambda_{0,\mathrm{tr}}(\Phi) = \sup_{Q\in\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}}_{0,\mathrm{tr}})} \big\{\Phi(Q) - I^{\mathrm{que}}_{\mathrm{tr}}(Q)\big\}. \tag{2.22}$$

Furthermore, for $\mathscr{W}$-a.e. X, the family $\mathscr{L}([R_N]_{\mathrm{tr}} \mid X)$, $N\in\mathbb{N}$, satisfies the LDP on $\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}}_{0,\mathrm{tr}})$ with deterministic rate function $I^{\mathrm{que}}_{\mathrm{tr}}$.

Note that the family of truncation operators $[\cdot]_{\mathrm{tr}}$ forms a projective system as the truncation level tr increases. Hence we immediately get from Proposition 2.4 and the Dawson-Gärtner projective limit LDP (see [7, Theorem 4.6.1]) that the family $\mathscr{L}(R_N \mid X)$, $N \in \mathbb{N}$, satisfies the LDP with rate function $Q \mapsto \sup_{\mathrm{tr}\in\mathbb{N}} I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}})$. Furthermore, since the projection can start at any initial level of truncation, we also know that the rate function is given by $Q \mapsto \sup_{\mathrm{tr}\geq n} I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}})$ for any $n\in\mathbb{N}$. Thus, Proposition 2.4 in fact implies that the rate function is given by

$$\tilde{I}^{\mathrm{que}}(Q) = \limsup_{\mathrm{tr}\to\infty} I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}}). \tag{2.23}$$

At this point, it remains to prove that $\tilde{I}^{\mathrm{que}}$ from (2.23) actually equals I^{que} from (1.13) and has the form claimed in (1.15).

This is achieved via the following proposition, note that (2.24) is the continuous analogue of [2, Lemma A.1].

Proposition 2.5. (1) *For $Q \in \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$,*

$$\lim_{\mathrm{tr}\to\infty} I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}}) = H(Q \mid Q_{\rho,\mathscr{W}}) + (\alpha-1) m_Q H(\Psi_Q \mid \mathscr{W}). \tag{2.24}$$

(2) *For $Q\in\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with $m_Q=\infty$ and $H(Q \mid Q_{\rho,\mathscr{W}})<\infty$ there exists a sequence $(\widetilde{Q}_{\mathrm{tr}})_{\mathrm{tr}\in\mathbb{N}}$ in $\mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$ such that* $\text{w-lim}_{\mathrm{tr}\to\infty}\widetilde{Q}_{\mathrm{tr}} = Q$ *and*

$$\tilde{I}^{\mathrm{que}}(\widetilde{Q}_{\mathrm{tr}}) \leq I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}}) + o(1), \qquad \mathrm{tr}\to\infty. \tag{2.25}$$

Proposition 2.5 (1) implies that for $Q\in\mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$ the lim sup in (2.23) is a limit, i.e., it implies (1.13) on $\mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$ and also (1.15).

To prove (1.13) for $Q\in\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with $m_Q=\infty$ and $H(Q \mid Q_{\rho,\mathscr{W}})<\infty$, consider $\widetilde{Q}_{\mathrm{tr}}$ as in Proposition 2.5 (2). Then

$$\tilde{I}^{\mathrm{que}}(Q) \leq \liminf_{\mathrm{tr}\to\infty} \tilde{I}^{\mathrm{que}}(\widetilde{Q}_{\mathrm{tr}}) \leq \liminf_{\mathrm{tr}\to\infty} I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}}), \tag{2.26}$$

where the first inequality uses that $\tilde{I}^{\mathrm{que}}$ is lower semi-continuous (being a rate function by the Dawson-Gärtner projective limit LDP), and the second inequality is a consequence of (2.25). For $Q\in\mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with $H(Q \mid Q_{\rho,\mathscr{W}})=\infty$ we have

$$\liminf_{\mathrm{tr}\to\infty} I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}}) \geq \liminf_{\mathrm{tr}\to\infty} H([Q]_{\mathrm{tr}} \mid [Q_{\rho,\mathscr{W}}]_{\mathrm{tr}}) = H(Q \mid Q_{\rho,\mathscr{W}}) = \infty, \tag{2.27}$$

i.e., also in this case the lim sup in (2.23) is a limit and (1.13) holds.

It remains to prove the properties of $I^{\rm que}$ claimed in Theorem 1.2: lower semi-continuity of $I^{\rm que} = \tilde{I}^{\rm que}$ follows from the representation via the Dawson-Gärtner projective limit LDP in (2.23); compactness of the level sets of $I^{\rm que}$ and the fact that $Q_{\rho,\mathscr{W}}$ is the unique zero of $Q \mapsto I^{\rm que}(Q)$ are inherited from the corresponding properties of $I^{\rm ann}$ because $I^{\rm que} \le I^{\rm ann}$; affineness of $Q \mapsto I^{\rm que}(Q)$ can be checked as in [2, Proof of Theorem 1.3]. □

Remark. Theorem 1.2 together with Varadhan's lemma implies that

$$\Lambda(\Phi) = \sup_{Q\in\mathcal{P}^{\rm inv,fin}(F^{\mathbb{N}})} \big\{\Phi(Q) - I^{\rm que}(Q)\big\}, \qquad \Phi\in\mathscr{C}, \tag{2.28}$$

and identifies $I^{\rm que}(Q)$ as the generalised convex transform

$$I^{\rm que}(Q) = \sup_{\Phi\in\mathscr{C}} \big\{\Phi(Q) - \Lambda(\Phi)\big\}, \qquad Q\in\mathcal{P}^{\rm inv}(F^{\mathbb{N}}) \tag{2.29}$$

(see [7, Theorems 4.4.2 and 4.4.10]). The supremum in (2.29) can also be taken over all continuous bounded functions on $\mathcal{P}^{\rm inv}(F^{\mathbb{N}})$.

2.3 Continuity of the empirical process under coarse-graining

Before embarking on the proof of Propositions 2.3–2.5 in Section 3, we state and prove two approximation lemmas (Lemmas 2.6–2.7 below) that will be needed along the way.

For $N\in\mathbb{N}$, $0=t_0<t_1<\cdots<t_N$ and $\varphi\in C([0,\infty))$, let $y_\varphi = (y_\varphi^{(i)})_{i\in\mathbb{N}}$ with

$$y_\varphi^{(i)} = \Big(t_i - t_{i-1}, \big(\varphi((t_{i-1}+s)\wedge t_i) - \varphi(t_{i-1})\big)_{s\ge 0}\Big) \in F, \qquad i=1,\dots,N, \tag{2.30}$$

and define

$$R_{N;t_1,\dots,t_N}(\varphi) = \frac{1}{N}\sum_{i=0}^{N-1} \delta_{\widetilde{\theta}^i y_\varphi^{N\text{-per}}} \in \mathcal{P}^{\rm inv}(F^{\mathbb{N}}). \tag{2.31}$$

We need a Skorohod-type distance d_S on paths, which is defined in Appendix A.

Lemma 2.6. *Let $i,j\in\mathbb{N}$, $i\le j$, and $t,t'\in(0,\infty)$, $t<t'$, be such that $(i-1)h<t\le ih$, $(j-1)h<t'\le jh$. Then, for any $\varphi\in C([0,\infty))$ and $k\in\mathbb{N}$,*

$$\begin{aligned} &d_S\big(\varphi((ih+\cdot)\wedge jh), \varphi((t+\cdot)\wedge t')\big)\\ &\le \log\tfrac{k+1}{k} + 2\sup_{(i-1)h\le s\le (i+k)h} |\varphi(s)-\varphi((i-1)h)| + 2\sup_{(j-1)h\le s\le jh} |\varphi(s)-\varphi((j-1)h)|. \end{aligned} \tag{2.32}$$

The same bound holds for $d_S([\varphi((ih+\cdot)\wedge jh)]_{\rm tr}, [\varphi((t+\cdot)\wedge t')]_{\rm tr})$ for any truncation length $\mathrm{tr}>0$.

Proof. Without loss of generality we may assume that $j\ge i+k$ (otherwise, employ the trivial time transform $\lambda(s)=s$ and estimate the left-hand side of (2.32) by the second term in the right-hand side of (2.32)), and use the time transformation

$$\lambda(s) = \begin{cases} s\,\frac{(i+k)h-t}{kh} & \text{if } s<kh,\\ s+ih-t & \text{if } s\ge kh. \end{cases} \tag{2.33}$$

In that case $\lambda(s)+t = s+ih$ for $s\ge kh$ and $\gamma(\lambda) = |\log[((i+k)h-t)/kh]| \le \log\frac{k+1}{k}$. The same argument applies to the truncated paths $[\varphi((ih+\cdot)\wedge jh)]_{\rm tr}$ and $[\varphi((t+\cdot)\wedge t')]_{\rm tr}$ (in fact, we can drop the third term in the right-hand side of (2.32) when $(j-1)h>\mathrm{tr}$). □

Lemma 2.7. *Let $\varphi \in C([0,\infty))$, $h > 0$, $N \in \mathbb{N}$ and $t_0 = 0 < t_1 < \cdots < t_N$. Let $\ell \in \mathbb{N}$, and let $g\colon \widehat{F}^\ell \to \mathbb{R}$ be bounded Lipschitz with Lipschitz constant C_g. Then, for $k \in \mathbb{N}$ with $k \geq \ell$,*

$$\begin{aligned} N\Big| \int_{\widehat{F}^\ell} g \, d\pi_\ell R_{N;t_1,\dots,t_N}(\varphi) - \int_{\widehat{F}^\ell} g \, d\pi_\ell R_{N;\lceil t_1\rceil_h,\dots,\lceil t_N\rceil_h}(\varphi)\Big| \\ \leq 4\ell\|g\|_\infty + C_g \ell N\big(2h + \log \tfrac{k+1}{k}\big) + 4C_g\ell \sum_{i=1}^N \sup_{\lceil t_i\rceil_h - h \leq s \leq \lceil t_i\rceil_h + kh} \big|\varphi(s) - \varphi(\lceil t_i\rceil_h - h)\big|, \end{aligned} \tag{2.34}$$

where $\pi_\ell\colon \widehat{F}^{\mathbb{N}} \to \widehat{F}^\ell$ denotes the projection onto the first ℓ coordinates. The same bound holds for the truncated versions $[R_{N;t_1,\dots,t_N}(\varphi)]_{\mathrm{tr}}$ and $[R_{N;\lceil t_1\rceil_h,\dots,\lceil t_N\rceil_h}(\varphi)]_{\mathrm{tr}}$ for any truncation length $\mathrm{tr} > 0$.

Proof. For $i = 1,\dots,N$, recall $y_\varphi^{(i)}$ from (2.30), i.e., $y_\varphi^{(i)}$ is the i-th word obtained by cutting the continuous path φ along the time points $t_1,\dots,t_n$, and let

$$\tilde{y}_\varphi^{(i,h)} = \Big(\lceil t_i\rceil_h - \lceil t_{i-1}\rceil_h, \big(\varphi((\lceil t_{i-1}\rceil_h + s) \wedge \lceil t_i\rceil_h) - \varphi(\lceil t_{i-1}\rceil_h)\big)_{s\geq 0}\Big), \tag{2.35}$$

be the analogous quantity when the h-discretised time points $\lceil t_1\rceil_h,\dots,\lceil t_N\rceil$ are used. By Lemma 2.6 we have

$$\begin{aligned} d_F\big(y_\varphi^{(i)}, \tilde{y}_\varphi^{(i,h)}\big) \leq \big(2h + \log \tfrac{k+1}{k}\big) + 2 \sup_{\lceil t_{i-1}\rceil_h - h \leq s \leq \lceil t_{i-1}\rceil_h + kh} \big|\varphi(s) - \varphi(\lceil t_{i-1}\rceil_h - h)\big| \\ + 2 \sup_{\lceil t_i\rceil_h - h \leq s \leq \lceil t_i\rceil_h} \big|\varphi(s) - \varphi(\lceil t_i\rceil_h - h)\big|. \end{aligned} \tag{2.36}$$

Writing $\tilde{y}^{(h)} = (\tilde{y}^{(i,h)})_{i\in\mathbb{N}}$ and putting, similarly as in (2.31),

$$R_{N;\lceil t_1\rceil_h,\dots,\lceil t_N\rceil_h}(\varphi) = \frac{1}{N}\sum_{i=0}^{N-1} \delta_{\widetilde{\theta}^i(\tilde{y}^{(h)})^{N\text{-per}}}, \tag{2.37}$$

we see that the claim follows from (2.16) in combination with Lemma 2.6. Note that possible boundary effects due to the periodisation are estimated by the term $4\ell\|g\|_\infty$. The observation about the truncated versions of the empirical process follow analogously from Lemma 2.6. □

3 Proof of Propositions 2.3–2.5

3.1 Proof of Proposition 2.3

Proof. The proof comes in 3 Steps.

Step 1. A.s. exponential tightness of the family $\mathscr{L}(R_N \mid X)$, $N \in \mathbb{N}$, is standard, because the family of unconditional distributions $\mathscr{L}(R_N)$ satisfies the LDP with a rate function that has compact level sets. Indeed, let $M > 0$, and pick a compact set $K \subset \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ such that $\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(R_N \notin K) \leq -2M$. Then $\mathbb{P}(\mathbb{P}(R_N \notin K \mid X) > e^{-MN}) \leq e^{MN}\mathbb{E}[\mathbb{P}(R_N \notin K \mid X)] \leq \exp(MN - 2MN + o(N))$, which is summable in N. Hence we have $\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}(R_N \notin K \mid X) \leq -M$ a.s. by the Borel-Cantelli lemma. The same argument applies to $[R_N]_{\mathrm{tr}}$ (alternatively, use the fact that $[\cdot]_{\mathrm{tr}}$ is a continuous map).

Step 2a. We next verify that the limits in (2.18) exist. In Step 2a we consider the case $\mathrm{supp}(\rho)=[0,\infty)$, in Step 2b the case $\mathrm{supp}(\rho)=[s_*,\infty)$ with $s_*>0$.

Let $\mathrm{tr}\in\mathbb{N}$ and $h=2^{-n}$. Let $Y^{(i,h)}=(\lceil T_i\rceil_h-\lceil T_{i-1}\rceil_h,(X_{(s+\lceil T_{i-1}\rceil_h)\wedge\lceil T_i\rceil_h}-X_{\lceil T_{i-1}\rceil_h})_{s\geq 0})\in\widehat{F}$ be the h-discretised i-th word, and let

$$R_{N,h}=\frac{1}{N}\sum_{i=0}^{N-1}\delta_{\widetilde{\theta}^i(Y^{(h)})^{N\text{-per}}} \tag{3.1}$$

be the h-discretised empirical process, where $Y^{(h)}=(Y^{(i,h)})_{i\in\mathbb{N}}$. Put $\ell=\ell_1\vee\cdots\vee\ell_m$, $C_g=C_{g_1}\vee\cdots\vee C_{g_m}$. Let

$$D_{j,h}=\sup_{(j-1)h\leq s\leq jh}|X_s-X_{jh}|,\qquad A_{\varepsilon,k,h}(N)=\left\{\sum_{i=1}^{N}\sum_{j=0}^{k}D_{\lceil T_i/h\rceil+j,h}\leq N\varepsilon\right\}. \tag{3.2}$$

By Lemma 2.7, on the event $A_{\varepsilon,k,h}(N)$ we have

$$N\big|\Phi([R_N]_{\mathrm{tr}})-\Phi([R_{N,h}]_{\mathrm{tr}})\big|\leq 4\ell\|\Phi\|_\infty+NC_g\ell m\Big(2h+\log\tfrac{k+1}{k}+4\varepsilon\Big), \tag{3.3}$$

and hence

$$\begin{aligned}\mathbb{E}\big[e^{N\Phi([R_N]_{\mathrm{tr}})}\big|X\big]\leq{}&\exp\big[NC_g\ell m\big(2h+\log\tfrac{k+1}{k}+4\varepsilon\big)+4\ell\|\Phi\|_\infty\big]\,\mathbb{E}\big[e^{N\Phi([R_{N,h}]_{\mathrm{tr}})}\big|X\big]\\&+e^{N\|\Phi\|_\infty}\mathbb{P}\big(A_{\varepsilon,k,h}(N)^c\mid X\big),\end{aligned} \tag{3.4}$$

For $\lambda>0$, estimate

$$\mathbb{P}([A_{\varepsilon,k,h}(N)]^c|X)\leq e^{-N\lambda\varepsilon}\,\mathbb{E}\left[\exp\left[\lambda\sum_{i=1}^{N}\sum_{m=0}^{k}D_{\lceil T_i/h\rceil+m,h}\right]\,\Big|\,X\right], \tag{3.5}$$

so that, by Lemma 3.2 in Step 4 below,

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big([A_{\varepsilon,k,h}(N)]^c\mid X\big)\leq-\varepsilon\lambda+\frac{1}{2}\log\chi\big(2k\lambda\sqrt{h}\big). \tag{3.6}$$

Since $\lim_{u\downarrow 0}\chi(u)=1$, we have, for all $\varepsilon>0$ and $k\in\mathbb{N}$,

$$\limsup_{h\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big([A_{\varepsilon,k,h}(N)]^c\mid X\big)=-\infty\quad\text{a.s.} \tag{3.7}$$

(pick $\lambda=\lambda(h)$ in (3.6) in such a way that $\lambda\to\infty$ and $\lambda\sqrt{h}\to 0$).

Next, observe that

$$\begin{aligned}\mathbb{E}\big[e^{N\Phi([R_N]_{\mathrm{tr}})}\mid X\big]={}&\int\cdots\int_{0<t_1<\cdots<t_N}\bar\rho(t_1)dt_1\,\bar\rho(t_2-t_1)dt_2\times\cdots\times\bar\rho(t_N-t_{N-1})dt_N\\&\times\exp\big[N\Phi\big([R_{N;t_1,\dots,t_N}(X)]_{\mathrm{tr}}\big)\big],\end{aligned} \tag{3.8}$$

$$\mathbb{E}\big[e^{N\Phi([R_{N,h}]_{\mathrm{tr}})}\mid X\big]=\sum_{1\leq j_1\leq\cdots\leq j_N}w_h(j_1,\dots,j_N)\exp\big[N\Phi\big([R_{N;hj_1,\dots,hj_N}(X)]_{\mathrm{tr}}\big)\big], \tag{3.9}$$

where

$$w_h(j_1,\dots,j_N) = \int\cdots\int_{0<t_1<\cdots<t_N} \bar\rho(t_1)dt_1\,\bar\rho(t_2-t_1)dt_2\times\cdots\times\bar\rho(t_N-t_{N-1})dt_N \times\prod_{k=1}^N 1_{(h(j_k-1),hj_k]}(t_k). \tag{3.10}$$

The idea is to replace the right-hand side of (3.10) by $\prod_{k=1}^N \lceil\rho\rceil_h(h(j_k-j_{k-1}))$, which is the corresponding weight for a discrete-time renewal process with waiting time distribution $\lceil\rho\rceil_h$. The rigorous implementation of this idea requires some care, since the coarse graining can produce "empty" words.

For $\underline{j}=(j_1,\dots,j_N)$ appearing in the sum in (3.9), let $R(\underline{j})=\#\{1\le i\le N\colon j_i=j_{i-1}\}$ be the total number of repeated values and $\underline{\hat\jmath}=(\hat\jmath_1,\dots,\hat\jmath_M)$ with $M=M(\underline{j})=N-R(\underline{j})$, $1\le\hat\jmath_1<\cdots<\hat\jmath_M$, the unique elements of $\underline{j}$. Note that any given $\underline{\hat\jmath}$ with $M=\lceil(1-\varepsilon)N\rceil$ can be obtained in this way from at most $\binom{N}{\lceil\varepsilon N\rceil}$ different $\underline{j}$'s.

In the following, we write $\eta(h)=\eta_n$ and $A(h)=A_n$ with η_n and A_n from (2.14) when $h=2^{-n}$. Let us parse through the right-hand side of (3.10) successively for $k=N,N-1,\dots,1$. When $j_k=j_{k-1}$, we integrate t_k out over $(h(j_k-1),hj_k]$ and estimate the (multiplicative) contribution of this integral from above by 1. When $j_k>j_{k-1}$, we replace $\bar\rho(t_k-t_{k-1})$ by $\bar\rho(t_k-hj_{k-1})$ and integrate t_k out over $(h(j_k-1),hj_k]$. For $h(j_k-j_{k-1})\in A(h)$ we can estimate the contribution of this integral from above by $e^{\eta(h)}\lceil\rho\rceil_h(h(j_k-j_{k-1})))$ by using (2.14), while for $h(j_k-j_{k-1})\notin A(h)$ we can estimate it by $e^{\eta_0}\lceil\rho\rceil_h(h(j_k-j_{k-1}))$ by using (2.15) with $s_*=0$. Thus, for $\underline{j}$ with $R(\underline{j})\le\varepsilon N$ and $\#\{1\le i<N\colon h(j_i-j_{i-1})\notin A(h)\}\le\varepsilon N$, we have

$$w_h(\underline{j})\le e^{\varepsilon\eta_0 N}e^{\eta(h)N}\prod_{i=1}^M\lceil\rho\rceil_h\big(h(\hat\jmath_i-\hat\jmath_{i-1})\big) \tag{3.11}$$

with $M=N-R(\underline{j})$. Furthermore,

$$\Big|N\Phi\big([R_{N;hj_1,\dots,hj_N}(X)]_{\rm tr}\big)-M\Phi\big([R_{M;h\hat\jmath_1,\dots,h\hat\jmath_M}(X)]_{\rm tr}\big)\Big|\le(N-M)\ell\|\Phi\|_\infty\le\varepsilon N\ell\|\Phi\|_\infty. \tag{3.12}$$

Combining (3.9–3.12), we find

$$\begin{aligned}
&\mathbb{E}\big[e^{N\Phi([R_{N,h}]_{\rm tr})}\mid X\big]\\
&\le e^{N\|\Phi\|_\infty}\Big\{\mathbb{P}\Big(R(\lceil T_1\rceil_h,\dots,\lceil T_N\rceil_h)\ge\varepsilon N\,\Big|\,X\Big)\\
&\qquad+\mathbb{P}\Big(\#\{1\le i<N\colon\lceil T_i\rceil_h-\lceil T_{i-1}\rceil_h\notin A(h)\}\ge\varepsilon N\,\Big|\,X\Big)\Big\}\\
&\quad+e^{[\varepsilon\eta_0+\eta(h)]N}\binom{N}{\varepsilon N}\sum_{M=\lceil(1-\varepsilon)N\rceil}^N\ \sum_{1\le\hat\jmath_1<\cdots<\hat\jmath_M}e^{M\Phi\big([R_{M;h\hat\jmath_1,\dots,h\hat\jmath_M}(X)]_{\rm tr}\big)}\prod_{k=1}^M\lceil\rho\rceil_h(h(\hat\jmath_k-\hat\jmath_{k-1})).
\end{aligned} \tag{3.13}$$

But

$$\sum_{1\le\hat\jmath_1<\cdots<\hat\jmath_M}e^{M\Phi\big([R_{M;h\hat\jmath_1,\dots,h\hat\jmath_M}(X)]_{\rm tr}\big)}\prod_{k=1}^M\lceil\rho\rceil_h(h(\hat\jmath_k-\hat\jmath_{k-1}))=\mathbb{E}_{\lceil\rho\rceil_h}\big[e^{M\Phi([R_M]_{\rm tr})}\mid X\big], \tag{3.14}$$

where $\mathbb{E}_{\lceil\rho\rceil_h}$ denotes expectation w.r.t. the reference measure $Q_{\lceil\rho\rceil_h,\mathscr{W}}$, and so we can apply Corollary 2.2 and Varadhan's lemma to obtain

$$\lim_{M\to\infty}\frac{1}{M}\log\mathbb{E}_{\lceil\rho\rceil_h}\big[e^{M\Phi([R_M]_{\rm tr})}\mid X\big]=\sup_{Q\in\mathcal{P}^{\rm inv,fin}(\widetilde{E_h}^{\mathbb{N}})}\big\{\Phi(Q)-I^{\rm que}_{h,\rm tr}(Q)\big\}. \tag{3.15}$$

By elementary large deviation estimates for binomials we have, for any $\varepsilon>0$,

$$\limsup_{h\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(R(\lceil T_1\rceil_h,\dots,\lceil T_N\rceil_h)\geq\varepsilon N\,\big|\,X\big)=-\infty, \tag{3.16}$$

$$\limsup_{h\downarrow 0}\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\Big(\#\big\{1\leq i<N\colon\ \lceil T_i\rceil_h-\lceil T_{i-1}\rceil_h\notin A(h)\big\}\geq\varepsilon N\,\Big|\,X\Big)=-\infty. \tag{3.17}$$

(Note that the events in (3.16–3.17) are independent of X.) Combining (3.4), (3.13) and (3.15), and noting that $\lim_{N\to\infty}\frac{1}{N}\log\binom{N}{\varepsilon N}=-\varepsilon\log\varepsilon-(1-\varepsilon)\log(1-\varepsilon)$, we find

$$\begin{aligned}
&\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{E}\big[e^{N\Phi([R_N]_{\rm tr})}\mid X\big]\\
&\quad\leq\Big\{\sup_{Q\in\mathcal{P}^{\rm inv,fin}((\widetilde{E_h})^{\mathbb{N}})}\big\{\Phi(Q)-I^{\rm que}_{h,\rm tr}(Q)\big\}\\
&\qquad\quad+C_g\ell m\big(2h+\log\tfrac{k+1}{k}+4\varepsilon\big)+\varepsilon\eta_0+\eta(h)+\varepsilon\log\tfrac{1}{\varepsilon}+(1-\varepsilon)\log\tfrac{1}{1-\varepsilon}\Big\}\\
&\quad\vee\Big(\|\Phi\|_\infty+\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\big(A_{\varepsilon,k,h}(N)^c\mid X\big)\Big\}\\
&\quad\vee\Big\{\|\Phi\|_\infty+\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\Big(R(\lceil T_1\rceil_h,\dots,\lceil T_N\rceil_h)\geq\varepsilon N\,\Big|\,X\Big)\Big\}\\
&\quad\vee\Big\{\|\Phi\|_\infty+\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}\Big(\#\{1\leq i<N\colon\ \lceil T_i\rceil_h-\lceil T_{i-1}\rceil_h\notin A(h)\}\geq\varepsilon N\,\Big|\,X\Big)\Big\},
\end{aligned} \tag{3.18}$$

and hence

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{E}\big[e^{N\Phi([R_N]_{\rm tr})}\big|X\big]\leq\liminf_{h\downarrow 0}\sup_{Q\in\mathcal{P}^{\rm inv,fin}(\widetilde{E_h}^{\mathbb{N}})}\big\{\Phi(Q)-I^{\rm que}_{h,\rm tr}(Q)\big\} \tag{3.19}$$

(let $h\downarrow 0$ along a suitable subsequence, followed by $\varepsilon\downarrow 0$ and $k\to\infty$, and use (3.7) and (3.16–3.17)).

Analogous arguments yield

$$\liminf_{N\to\infty}\frac{1}{N}\log\mathbb{E}\big[e^{N\Phi([R_N]_{\rm tr})}\mid X\big]\geq\limsup_{h\downarrow 0}\sup_{Q\in\mathcal{P}^{\rm inv,fin}((\widetilde{E_h})^{\mathbb{N}})}\big\{\Phi(Q)-I^{\rm que}_{h,\rm tr}(Q)\big\}. \tag{3.20}$$

Indeed, we can simply restrict the sum in (3.9) to $\underline{j}$'s with $j_1<\cdots<j_N$, so that the approximation argument is in fact a little easier because we need not pass to the $\underline{\hat{\jmath}}$'s.

Finally, combine (3.19–3.20) to obtain (2.18).

Step 2b. Next we consider the case $\mathrm{supp}(\rho)=[s_*,\infty)$ with $s_*>0$ and indicate the changes compared to Step 2a. To some extent this case is easier than the case $s_*=0$, since for coarse-graining level $h<s_*$ no "empty" word can appear in the coarse-graining scheme. On

the other hand, when implementing a replacement similar to (3.11), it can happen that an integral $\int \bar\rho(t_k - t_{k-1}) 1_{(h(j_k-1), hj_k]}(t)\, dt_k$ gets mapped to $\lceil\rho\rceil_h(h(j_k - j_{k-1})) = 0$ even though the true contribution of that integral to (3.9) is strictly positive (namely, when $h(j_k - j_{k-1}) \le s_* \le h(j_k - j_{k-1} + 1)$). The idea to remedy this problem is to replace $\lceil\rho\rceil_h(h(j_k - j_{k-1}))$ by a sum of "neighbouring" weights of $\lceil\rho\rceil_h$ and to suitably control the overcounting incurred by this replacement. The details are as follows.

Fix $h > 0$ and $s_{*,h} = \lceil s_*\rceil_h$. For $N \in N$, consider $\underline{j} = (j_1, \dots, j_N)$ as appearing in the sum in (3.9). We say that $k \in \{1, \dots, N\}$ is "problematic" when $h(j_k - j_{k-1}) \in \{s_{*,h} - 1, s_{*,h}, s_{*,h} + 1\}$, and "relaxable" when $j_k - j_{k-1} \ge 2$ and

$$\max_{m=-1,0,1} \left| \log \frac{\lceil\rho\rceil_h(h(j_k - j_{k-1} + m))}{\lceil\rho\rceil_h(h(j_k - j_{k-1}))} \right| \le 2. \tag{3.21}$$

Write $K_{\mathrm{pro}}(\underline{j}) = \{1 \le k \le N\colon\ k \text{ problematic}\}$ and $K_{\mathrm{rel}}(\underline{j}) = \{1 \le k \le N\colon\ k \text{ relaxable}\}$. Try to construct an injection $f_{\mathrm{rel},\underline{j}}\colon K_{\mathrm{pro}} \to K_{\mathrm{rel}}$ with the property $f_{\mathrm{rel},\underline{j}}(k) > k$ as follows:

> Start with an empty "stack" s. For $k = 1, \dots, N$ successively: when k is problematic, push k on s; when k is relaxable and s is not empty, pop the top element, say k', from s and put $f_{\mathrm{rel},\underline{j}}(k') = k$; when k is neither problematic nor relaxable, proceed with the next k.

We say that $\underline{j}$ is "good" when the above procedure terminates with an empty stack (in particular, $f_{\mathrm{rel},\underline{j}}(k')$ is defined for all $k' \in K_{\mathrm{pro}}$) and

$$\sum_{k \in K_{\mathrm{pro}}} \big(f_{\mathrm{rel},\underline{j}}(k) - k \big) \le \varepsilon N \tag{3.22}$$

(in particular, $\#K_{\mathrm{pro}}(\underline{j}) \le \varepsilon N$), and also $\#\{1 \le k \le N\colon\ j_k - j_{k-1} \notin A(h)\} \le \varepsilon N$. For a given good $\underline{j}$, consider the set of all $\underline{\tilde{j}} = (\tilde{j}_1, \dots, \tilde{j}_N)$ obtainable by setting

$$\tilde{j}_k = j_k + \Delta_k,\ \tilde{j}_{f_{\mathrm{rel},\underline{j}}(k)} = j_{f_{\mathrm{rel},\underline{j}}(k)} - \Delta_k \quad \text{with } \Delta_k \in \{-1, 0, 1\} \quad \text{for } k \in K_{\mathrm{pro}}, \tag{3.23}$$

and $\tilde{j}_k = j_k$ for $k \notin (K_{\mathrm{pro}} \cup f_{\mathrm{rel},\underline{j}}(K_{\mathrm{pro}}))$. Note that a given good $\underline{j}$ corresponds to at most $3^{\varepsilon N}$ different $\underline{\tilde{j}}$'s and that, for any such $\underline{\tilde{j}}$,

$$\begin{aligned} &\Big| N\Phi\big([R_{N;hj_1,\dots,hj_N}(X)]_{\mathrm{tr}}\big) - N\Phi\big([R_{N;h\tilde{j}_1,\dots,h\tilde{j}_M}(X)]_{\mathrm{tr}}\big) \Big| \\ &\qquad \le \ell \|\Phi\|_\infty \sum_{k \in K_{\mathrm{pro}}} \big(f_{\mathrm{rel},\underline{j}}(k) - k \big) \le \varepsilon N \ell \|\Phi\|_\infty. \end{aligned} \tag{3.24}$$

With $w_h(j_1, \dots, j_N)$ defined in (3.10), we now see that (analogously to the argument prior to (3.11)) for any good $\underline{j}$,

$$w_h(\underline{j}) \le e^{\varepsilon \eta_0 N} e^{\eta(h) N} 2^{\varepsilon N} \sum_{\underline{\tilde{j}} \text{ corresp. to } \underline{j}} \prod_{i=1}^N \lceil\rho\rceil_h\big(h(\tilde{j}_i - \tilde{j}_{i-1})\big). \tag{3.25}$$

Moreover, we have

$$\limsup_{h \downarrow 0} \limsup_{N \to \infty} \frac{1}{N} \log \mathbb{P}\big((\lceil T_1\rceil_h, \dots, \lceil T_N\rceil_h) \text{ not good} \,\big|\, X \big) = -\infty. \tag{3.26}$$

To check (3.26), let S_k be the size of the stack s in the k-th step of the above construction when we use $j_k = \lceil T_k \rceil_h$, and note that $(\lceil T_1 \rceil_h, \dots, \lceil T_N \rceil_h)$ is good when $\sum_{k=1}^N S_k < \varepsilon N$. A comparison of $(S_k)_{k\in\mathbb{N}}$ with a (reflected) random walk on $\mathbb{N}_0$ that draws its steps from $\{0, \pm 1\}$, where $(+1)$-steps have a very small probability ($\leq \int_{s_*}^{s_*+2h} \bar\rho(t)\,dt$) and (-1)-steps have a very large probability ($\rho(A_h)$) when not from 0, shows that $\limsup_{h\downarrow h} \frac{1}{N} \log \mathbb{P}(\sum_{k=1}^N S_k \geq \varepsilon N) = -\infty$ for every $\varepsilon > 0$. We can then estimate similarly as in (3.18), to obtain (3.19) for the case $s_* > 0$ as well.

Analogous arguments also yield the lower bound in (3.20).

Step 3. We next verify that the limits in (2.20) exist. Note that

$$|\Phi(R_N) - \Phi([R_N]_{\mathrm{tr}})| \leq \|\Phi\|_\infty \frac{1}{N} \#\big\{\text{loops among the first } N \text{ loops that are longer than tr}\big\}, \tag{3.27}$$

which can be made arbitrarily small (also on the exponential scale, via a suitable annealing argument that uses that loop lengths are i.i.d.). A similar estimate holds for $|\Phi([R_N]_{\mathrm{tr}}) - \Phi([R_N]_{\mathrm{tr}'})|$ with $\mathrm{tr} < \mathrm{tr}'$. This shows that $\Lambda_{0,\mathrm{tr}}(\Phi)$ forms a Cauchy sequence as $\mathrm{tr} \to \infty$. □

Remark 3.1. *The arguments in Steps* 2a *and* 2b *can be combined to yield the same results when assumption* (1.2) *is relaxed to assumption* (1.24)*. Indeed, for a given coarse-graining level h,* (1.24) *gives rise to finitely many types of "problematic points" that can be handled similarly as in Step* 2b *(combined with arguments from Step* 2a *when $a_1 = 0$).*

Step 4. We close by deriving the estimate on Brownian increments over randomly drawn short time intervals that was used in (3.6) in Step 2. The intuitive idea is that even though there are arbitrarily large increments over short time intervals somewhere on the Brownian path, it is extremely unlikely to hit these when sampling along an independent renewal process. The proof employs a suitable annealing argument.

Recall $D_{j,h}$ from (3.2). For $h > 0$ fixed, the $D_{j,h}$'s are i.i.d. and equal in law to $\sqrt{h} D_{1,1} = \sqrt{h} \sup_{0\leq s\leq 1} |X_s|$ by Brownian scaling.

Lemma 3.2. *Let $T = (T_i)_{i\in\mathbb{N}}$ be a continuous-time renewal process with interarrival law ρ satisfying* $\mathrm{supp}(\rho) \subset [h, \infty)$*. For $\lambda \geq 0$ and $k \in \mathbb{N}_0$, define*

$$\xi(\lambda, h) = \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{E}\left[\exp\left[\lambda \sum_{i=1}^N \sum_{m=0}^k D_{\lceil T_i/h \rceil + m, h}\right] \,\Big|\, \sigma(D_{j,h}, j \in \mathbb{N})\right], \tag{3.28}$$

which is ≥ 0 and a.s. constant by Kolmogorov's 0-1*-law. Then*

$$\lim_{h\downarrow 0} \xi(\lambda, h) = 0 \qquad \forall\, \lambda \geq 0. \tag{3.29}$$

Proof. We consider only the case $k = 0$, the proof for $k \in \mathbb{N}$ being analogous. Abbreviate $\mathscr{G}_h = \sigma(D_{j,h}, j \in \mathbb{N})$, and let

$$\chi(u) = \mathbb{E}\Big[\exp\big[u \sup{}_{0\leq t\leq 1} |X_t|\big]\Big], \quad u \in \mathbb{R}. \tag{3.30}$$

Note that $\chi(\cdot)$ is finite and satisfies $\lim_{u\to 0} \chi(u) = 1$. We have

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[\exp\left[\lambda \sum_{i=1}^N D_{\lceil T_i/h \rceil, h}\right] \,\Big|\, \mathscr{G}_h\right]^2\right] &\leq \mathbb{E}\left[\exp\left[2\lambda \sum_{i=1}^N D_{\lceil T_i/h \rceil, h}\right]\right] \\ &= \mathbb{E}\big[\exp[2\lambda D_{1,h}]\big]^N = \chi\big(2\lambda\sqrt{h}\big)^N.\end{aligned} \tag{3.31}$$

Thus, for any $\epsilon > 0$,

$$\begin{aligned}&\mathbb{P}\left(\mathbb{E}\left[\exp\Big[\lambda\sum_{i=1}^{N} D_{\lceil T_i/h\rceil,h}\Big]\,\Big|\,\mathscr{G}_h\right]^2 \geq \big(\chi\big(2\lambda\sqrt{h}\big)+\epsilon\big)^N\right)\\&\leq\ \big(\chi\big(2\lambda\sqrt{h}\big)+\epsilon\big)^{-N}\mathbb{E}\left[\mathbb{E}\left[\exp\Big[\lambda\sum_{i=1}^{N} D_{\lceil T_i/h\rceil,h}\Big]\,\Big|\,\mathscr{G}_h\right]^2\right] \leq \left(\frac{\chi\big(2\lambda\sqrt{h}\big)}{\chi\big(2\lambda\sqrt{h}\big)+\epsilon}\right)^N,\end{aligned} \tag{3.32}$$

which is summable in N. The Borel-Cantelli lemma therefore yields

$$\limsup_{N\to\infty}\frac{1}{N}\log\mathbb{E}\left[\exp\Big[\lambda\sum_{i=1}^{N} D_{\lceil T_i/h\rceil,h}\Big]\,\Big|\,\mathscr{G}_h\right] \leq \frac{1}{2}\log\chi\big(2\lambda\sqrt{h}\big). \tag{3.33}$$

□

3.2 Proof of Proposition 2.4

Lemma 3.3. *For* $\mathrm{tr}\in\mathbb{N}$ *and* $Q\in\mathcal{P}^{\mathrm{inv}}(F_{0,\mathrm{tr}}^{\mathbb{N}})$,

$$I_{\mathrm{tr}}^{\mathrm{que}}(Q) = \lim_{\varepsilon\downarrow 0}\limsup_{h\downarrow 0}\inf\Big\{I_{h,\mathrm{tr}}^{\mathrm{que}}(Q')\colon\ Q'\in B_\varepsilon(Q)\cap\mathcal{P}^{\mathrm{inv}}((\widetilde{E}_{h,\mathrm{tr}})^{\mathbb{N}})\Big\}, \tag{3.34}$$

where $h\downarrow 0$ *along* 2^{-m}, $m\in\mathbb{N}$.

Note that after $\widetilde{E}_{h,\mathrm{tr}}$ is identified with a subset of $F_{0,\mathrm{tr}}$ (see (2.4)), (3.34) states that $I_{h,\mathrm{tr}}^{\mathrm{que}}$ converges to $I_{\mathrm{tr}}^{\mathrm{que}}$ as $h\downarrow 0$ in the sense of Gamma-convergence.

Proof of Lemma 3.3. Note that, when restricted to $\mathcal{P}^{\mathrm{inv}}(F_{0,\mathrm{tr}}^{\otimes\mathbb{N}})$,

$$\text{both } Q\mapsto m_Q \text{ and } Q\mapsto\Psi_Q \text{ are continuous} \tag{3.35}$$

(by dominated convergence), while this is not true when Q is allowed to vary over the whole of $\mathcal{P}^{\mathrm{inv}}(F^{\otimes\mathbb{N}})$. A more general statement is the following: if $\text{w-lim}_{n\to\infty}Q_n = Q$ and $\{\mathscr{L}_{Q_n}(\tau_1)\colon\ n\in\mathbb{N}\}$ are uniformly integrable, then $\lim_{n\to\infty}m_{Q_n} = m_Q$ and $\text{w-lim}_{n\to\infty}\Psi_{Q_n} = \Psi_Q$.

In the proof we use several properties of specific relative entropy derived in Appendix B. Let $Q\in\mathcal{P}^{\mathrm{inv}}(F_{0,\mathrm{tr}}^{\mathbb{N}})$, and abbreviate the right-hand side of (3.34) by $\widetilde{I}_{\mathrm{tr}}^{\mathrm{que}}(Q)$. Note that, by (3.35) and the lower semi-continuity of $\Psi\mapsto H(\Psi\mid\mathscr{W})$, the map

$$\mathcal{P}^{\mathrm{inv}}(F_{0,\mathrm{tr}}^{\mathbb{N}})\ni Q'\mapsto m_{Q'}H(\Psi_{Q'}\mid\mathscr{W}) \tag{3.36}$$

is lower semi-continuous. Hence, for any $\delta>0$, we have $m_{Q'}H(\Psi_{Q'}\mid\mathscr{W})\geq m_Q H(\Psi_Q\mid\mathscr{W})-\delta$ for all $Q'\in B_\varepsilon(Q)\cap\mathcal{P}^{\mathrm{inv}}(\widetilde{E}_{h,\mathrm{tr}}^{\mathbb{N}})$ when ε is sufficiently small (depending on δ). Combine this with (B.10) in Lemma B.2 in Appendix B, and note that $\text{w-lim}\,Q_{h,\mathrm{tr}} = Q_{\mathrm{tr}}$ as $h\downarrow 0$, to obtain $\widetilde{I}_{\mathrm{tr}}^{\mathrm{que}}(Q)\geq I_{\mathrm{tr}}^{\mathrm{que}}(Q)$.

For the reverse direction, we need to find $h_n>0$ with $\lim_{n\to\infty}h_n = 0$ and $Q_n'\in\mathcal{P}^{\mathrm{inv}}((\widetilde{E}_{h_n,\mathrm{tr}})^{\mathbb{N}})$ with $\text{w-lim}_{n\to\infty}Q_n' = Q$ such that $\liminf_{n\to\infty}I_{h_n,\mathrm{tr}}^{\mathrm{que}}(Q_n')\leq I_{\mathrm{tr}}^{\mathrm{que}}(Q)$. Here a complication stems from the fact that we must ensure that both parts of $I_{h_n,\mathrm{tr}}^{\mathrm{que}}(Q_n')$, namely, $H(Q_n'\mid Q_{\lceil\rho\rceil_{h_n},\mathscr{W},\mathrm{tr}})$ and $H(\Psi_{Q_n',h_n}\mid\mathscr{W})$, converge simultaneously. The proof is deferred to Lemma B.3 in Appendix B. □

We are now ready to give the proof of Proposition 2.4.

Proof. Fix $\mathrm{tr} \in \mathbb{N}$. Denote the right-hand side of (2.22) by $\tilde{\Lambda}_{\mathrm{tr}}(\Phi)$. Let $\Phi\colon \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}}) \to \mathbb{R}$ be of the form (2.17). For every $\delta > 0$ we can find a $Q^* \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}}_{0,\mathrm{tr}})$ such that $\Phi(Q^*) - I^{\mathrm{que}}_{\mathrm{tr}}(Q^*) \geq \tilde{\Lambda}_{\mathrm{tr}}(\Phi) - \delta$. For $\varepsilon > 0$ sufficiently small (depending on δ) we have $\left|\Phi(Q') - \Phi(Q^*)\right| \leq \delta$ for all $Q' \in B_\varepsilon(Q^*)$ and, by Lemma 3.3,

$$\liminf_{h\downarrow 0} \inf \left\{ I^{\mathrm{que}}_{h,\mathrm{tr}}(Q')\colon\ Q' \in B_\varepsilon(Q^*) \cap \mathcal{P}^{\mathrm{inv}}((\widetilde{E_{h,\mathrm{tr}}})^{\mathbb{N}}) \right\} \leq I^{\mathrm{que}}_{\mathrm{tr}}(Q^*) + \delta. \tag{3.37}$$

Thus

$$\liminf_{h\downarrow 0} \sup \left\{ \Phi(Q') - I^{\mathrm{que}}_{h,\mathrm{tr}}(Q')\colon\ Q' \in B_\varepsilon(Q^*) \cap \mathcal{P}^{\mathrm{inv}}((\widetilde{E_{h,\mathrm{tr}}})^{\mathbb{N}}) \right\} \geq \tilde{\Lambda}_{\mathrm{tr}}(\Phi) - 3\delta. \tag{3.38}$$

Let $\delta \downarrow 0$ to obtain $\liminf_{h\downarrow 0} \Lambda_{h,\mathrm{tr}}(\Phi) = \Lambda_{0,\mathrm{tr}}(\Phi) \geq \tilde{\Lambda}_{\mathrm{tr}}(\Phi)$.

For the reverse direction, pick for $h \in (0,1)$ a maximiser $Q^*_h \in \mathcal{P}^{\mathrm{inv}}((\widetilde{E_{h,\mathrm{tr}}})^{\mathbb{N}})$ of the variational expression appearing in the right-hand side of (2.19), i.e., $\Phi(Q^*_h) - I^{\mathrm{que}}_{h,\mathrm{tr}}(Q^*_h) = \Lambda_{h,\mathrm{tr}}(\Phi)$. This is possible because $\Phi - I^{\mathrm{que}}_{h,\mathrm{tr}}$ is upper semi-continuous and bounded from above, and $I^{\mathrm{que}}_{h,\mathrm{tr}}$ has compact level sets. We claim that

$$\text{the family } \left\{ Q^*_h : h \in (0,1) \right\} \subset \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}}) \text{ is tight.} \tag{3.39}$$

Assuming (3.39), we can choose a sequence $h(n) \downarrow 0$ such that

$$\begin{aligned} &\lim_{n\to\infty} \left[\Phi(Q^*_{h(n)}) - I^{\mathrm{que}}_{h(n),\mathrm{tr}}(Q^*_{h(n)}) \right] = \limsup_{h\downarrow 0} \Lambda_{h,\mathrm{tr}}(\Phi), \\ &\operatorname*{w-lim}_{n\to\infty} Q^*_{h(n)} = \widetilde{Q} \ \text{ for some } \widetilde{Q} \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}}). \end{aligned} \tag{3.40}$$

Then $\lim_{n\to\infty} \Phi(Q^*_{h(n)}) = \Phi(\widetilde{Q})$ because Φ is continuous, and $\liminf_{n\to\infty} I^{\mathrm{que}}_{h(n),\mathrm{tr}}(Q^*_{h(n)}) \geq I^{\mathrm{que}}_{\mathrm{tr}}(\widetilde{Q})$ by Lemma 3.3. Hence

$$\Lambda_{0,\mathrm{tr}}(\Phi) = \limsup_{h\downarrow 0} \Lambda_{h,\mathrm{tr}}(\Phi) = \lim_{n\to\infty} \left[\Phi(Q^*_{h(n)}) - I^{\mathrm{que}}_{h(n),\mathrm{tr}}(Q^*_{h(n)}) \right] \leq \Phi(\widetilde{Q}) - I^{\mathrm{que}}_{\mathrm{tr}}(\widetilde{Q}) \leq \tilde{\Lambda}_{\mathrm{tr}}(\Phi). \tag{3.41}$$

It remains to prove (3.39), which follows once we show that for each $N \in \mathbb{N}$ the family of projections $\pi_N(Q^*_h) \in \mathcal{P}^{\mathrm{inv}}(F^N)$, $h \in (0,1)$, is tight (because $F^{\mathbb{N}}$ carries the product topology; see Ethier and Kurtz [9, Chapter 3, Proposition 2.4]). Let $M = \|\Phi\|_\infty + 1$. Then necessarily $H(Q^*_h \mid [Q_{\lceil\rho\rceil_h,\mathscr{W}}]_{\mathrm{tr}}) \leq M$, and hence $h(\pi_N(Q^*_h) \mid \pi_N([Q_{\lceil\rho\rceil_h,\mathscr{W}}]_{\mathrm{tr}}) \leq NM$ for all $h \in (0,1)$. Since $\pi_N([Q_{\lceil\rho\rceil_h,\mathscr{W}}]_{\mathrm{tr}}) = ([q_{\lceil\rho\rceil_h,\mathscr{W}}]_{\mathrm{tr}})^{\otimes N}$ converges weakly to $\pi_N([Q_{\rho,\mathscr{W}}]_{\mathrm{tr}}) = ([q_{\rho,\mathscr{W}}]_{\mathrm{tr}})^{\otimes N}$ as $h \downarrow 0$, the family $\{\pi_N([Q_{\lceil\rho\rceil_h,\mathscr{W}}]_{\mathrm{tr}})\colon\ h \in (0,1)\}$ is tight, and so for any $\varepsilon > 0$ we can find a compact $\mathcal{C} \subset F^N$ such that $\pi_N([Q_{\lceil\rho\rceil_h,\mathscr{W}}]_{\mathrm{tr}})(\mathcal{C}^c) \leq \exp[-(NM + \log 2)/\varepsilon]$ uniformly in $h \in (0,1)$. By a standard entropy inequality (see (B.3) in Appendix B), for all $h \in (0,1)$ we have

$$\pi_N(Q^*_h)(\mathcal{C}^c) \leq \frac{\log 2 + h\big(\pi_N(Q^*_h) \mid \pi_N([Q_{\lceil\rho\rceil_h,\mathscr{W}}]_{\mathrm{tr}})\big)}{\log\Big(1 + \big(\pi_N([Q_{\lceil\rho\rceil_h,\mathscr{W}}]_{\mathrm{tr}})(\mathcal{C}^c)\big)^{-1}\Big)} \leq \frac{\log 2 + MN}{\log\big(1 + \exp[(NM + \log 2)/\varepsilon]\big)} \leq \varepsilon. \tag{3.42}$$

This proves the representation (2.22) of the limit $\Lambda_{0,\mathrm{tr}}(\Phi)$ from (2.18). From (2.18) and (2.22), plus the exponential tightness in Proposition 2.3, we obtain the LDP via Bryc's inverse of Varadhan's lemma. □

3.3 Proof of Proposition 2.5

3.3.1 Proof of part (1)

We first verify (2.24), i.e., for $Q \in \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$,

$$\begin{aligned}\lim_{\mathrm{tr}\to\infty} I^{\mathrm{que}}_{\mathrm{tr}}([Q]_{\mathrm{tr}}) &= \lim_{\mathrm{tr}\to\infty}\Big[H([Q]_{\mathrm{tr}} \mid [Q_{\rho,\mathscr{W}}]_{\mathrm{tr}}) + (\alpha-1)m_{[Q]_{\mathrm{tr}}} H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W})\Big]\\ &= H(Q \mid Q_{\rho,\mathscr{W}}) + (\alpha-1)m_Q H(\Psi_Q \mid \mathscr{W}). \end{aligned} \tag{3.43}$$

The proof comes in 5 Steps.

Step 1. Note that $\lim_{\mathrm{tr}\to\infty} H([Q]_{\mathrm{tr}} \mid [Q_{\rho,\mathscr{W}}]_{\mathrm{tr}}) = H(Q \mid Q_{\rho,\mathscr{W}})$ by the projective property of word truncations, $\lim_{\mathrm{tr}\to\infty} m_{[Q]_{\mathrm{tr}}} = m_Q < \infty$ by dominated convergence, and

$$\liminf_{\mathrm{tr}\to\infty} H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}) \geq H(\Psi_Q \mid \mathscr{W}) \tag{3.44}$$

by the lower semi-continuity of specific relative entropy together with w-$\lim_{\mathrm{tr}\to\infty} \Psi_{[Q]_{\mathrm{tr}}} = \Psi_Q$. Hence, to obtain (3.43) it remains to prove that

$$\limsup_{\mathrm{tr}\to\infty} H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}) \leq H(\Psi_Q \mid \mathscr{W}). \tag{3.45}$$

Step 2. To prove (3.45), we use coarse-graining. For every $h > 0$ we can identify $\widetilde{E_h}$ with $F_h \subset F$ (recall (2.3)). In order to represent $Q \in \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$ by a shift-invariant law on $(F_h)^{\mathbb{N}}$, we discretise the cut-points onto a *uniformly shifted* grid of width h, as follows. For $t \in \mathbb{R}$, $h > 0$ and $u \in [0,1)$, define (compare with Section 2.1.2)

$$\lceil t \rceil_{h,u} = \min\big\{(k+u)h \colon k \in \mathbb{Z}, (k+u)h \geq t\big\} \quad \big(= \lceil t - uh \rceil_h + uh\big). \tag{3.46}$$

Draw $Y = (Y^{(i)})_{i\in\mathbb{N}} = ((\tau_i, f_i))_{i\in\mathbb{N}}$ from law Q, and let U be an independent random variable with uniform distribution on $[0,1]$. Put $T_0 = 0$, $T_n = \tau_1 + \cdots + \tau_n$, $n \in \mathbb{N}$,

$$\tilde{T}_i = \lceil T_i \rceil_{h,U}, \quad i \in \mathbb{N}_0, \qquad \tilde{\tau}_i = \tilde{T}_i - \tilde{T}_{i-1},\ \tilde{f}_i = \big(\theta^{\tilde{T}_{i-1}}\kappa(Y)\big)(\,\cdot \wedge \tilde{\tau}_i), \quad i \in \mathbb{N}. \tag{3.47}$$

(Note that it may happen that $\tilde{\tau}_i = 0$. We can remedy this by allowing "empty words", i.e., by formally passing to $\widehat{F}$ as in Section 2.2.1.) Write $\lceil Q \rceil_h$ for the distribution of $\tilde{Y} = (\tilde{Y}^{(i)})_{i\in\mathbb{N}} = ((\tilde{\tau}_i, \tilde{f}_i))_{i\in\mathbb{N}}$ obtained in this way. We view $\lceil Q \rceil_h$ as an element of $\mathcal{P}^{\mathrm{inv,fin}}((F_h)^{\mathbb{N}})$. To check the shift-invariance of $\lceil Q \rceil_h$, note that by construction an initial part of length $S_1 = \tilde{T}_0 - T_0 = Uh$ of the content of the first word is removed (in a two-sided situation, this part would be added at the end of the zero-th word). The corresponding quantity for the second word is $S_2 = \tilde{T}_1 - T_1 = \lceil T_1 \rceil_{h,U} - T_1$. Observe that, for measurable $A \subset [0,h)$ and $B \subset [0,\infty)$,

$$\mathbb{P}(S_2 \in A, T_1 \in B) = \int_B \mathbb{P}(T_1 \in dt) \int_{[0,1]} du\, 1_A\big(\lceil t - uh \rceil_h - (t - uh)\big) = \frac{1}{h}\,\mathbb{P}(T_1 \in B)\,\lambda(A), \tag{3.48}$$

i.e., S_2 is distributed as Uh and independent of Y, and so $(\tilde{Y}^{(i+1)})_{i\in\mathbb{N}}$ again has law $\lceil Q \rceil_h$. This settles the shift-invariance. The key feature of the construction of $\lceil Q \rceil_h$ is that $\kappa(\tilde{Y}) = (\theta^{Uh}\kappa)(Y)$, so that

$$\Psi_{\lceil Q \rceil_h, h} = \Psi_Q, \tag{3.49}$$

and therefore

$$H(\Psi_{\lceil Q\rceil_h,h} \mid \mathscr{W}) = H(\Psi_Q \mid \mathscr{W}). \tag{3.50}$$

Thus, (3.50) gives us a coarse-grained version of the right-hand of (3.45).

Step 3. If tr is an integer multiple of h, then the coarse-graining $\lceil Q\rceil_h \in \mathcal{P}^{\mathrm{inv,fin}}((F_h)^{\mathbb{N}})$ of $Q \in \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$ defined in Step 2 commutes with the word length truncation $[\cdot]_{\mathrm{tr}}$, i.e., $[\lceil Q\rceil_h]_{\mathrm{tr}} = \lceil [Q]_{\mathrm{tr}}\rceil_h$. This is a deterministic property of the construction in (3.46). Indeed, fix $u \in [0,1)$ and h with $\mathrm{tr} = Mh$ for some $M \in \mathbb{N}$, consider $t_{i-1} < t_i$ with $t_i - t_{i-1} > \mathrm{tr}$ (so that in the un-coarse-grained truncation procedure the i-th loop length would be replaced by tr), let $k_{i-1}, k_i \in \mathbb{N}$ be such that $\lceil t_{i-1}\rceil_{h,u} = (k_{i-1}+u)h$ and $\lceil t_i\rceil_{h,u} = (k_i+u)h$. When we first truncate and then coarse-grain, the i-th point becomes $\lceil t_{i-1}+\mathrm{tr}\rceil_{h,u} = (k_{i-1}+M+u)h$. When we first coarse-grain and then truncate, the i-th point becomes $\lceil t_{i-1}\rceil_{h,u} + \big((\lceil t_i\rceil_{h,u} - \lceil t_{i-1}\rceil_{h,u}) \wedge Mh\big) = (k_{i-1}+u)h + Mh$, which is the same.

Step 4. Let $h = 2^{-M}$, define $\lceil Q\rceil_h \in \mathcal{P}^{\mathrm{inv,fin}}((F_h)^{\mathbb{N}})$ as in Step 2, and write $Q'_h = \lceil Q\rceil_h \circ \iota_h^{-1}$ for the same object considered as an element of $\mathcal{P}^{\mathrm{inv,fin}}((\widetilde{E_h})^{\mathbb{N}})$ (recall (2.1–2.4)). Write $\nu_h = \mathscr{L}\big((X_{\cdot\wedge h})\big)$ for the Wiener measure on E_h. Then $m_{Q'_h} = m_{\lceil Q\rceil_h}/h$ (the mean word length counted in h-letters), while

$$H(\Psi_{Q'_h} \mid \nu_h^{\otimes\mathbb{N}}) = H(\Psi_{\lceil Q\rceil_h,h} \mid \mathscr{W}), \tag{3.51}$$

by construction, and

$$\lceil [Q]_{\mathrm{tr}}\rceil_h = [\lceil Q\rceil_h]_{\mathrm{tr}} = [Q'_h]_{(\mathrm{tr}/h)} \circ \iota_h, \tag{3.52}$$

where the first equality follows from the commutation property in Step 3 and the second equality is a truncation of the words from Q'_h as elements of $\widetilde{E_h}$.

Step 5. Fix $\varepsilon > 0$ and let $\mathrm{tr}_0 = \mathrm{tr}_0(Q,\varepsilon)$ be so large that

$$\mathbb{E}_Q\big[\big(|Y^{(1)}| - \mathrm{tr}\big)_+\big] < \tfrac{1}{3}\varepsilon m_Q, \qquad \mathrm{tr} \geq \mathrm{tr}_0. \tag{3.53}$$

Then, for $0 < h < \frac{1}{24}\varepsilon m_Q$, we have

$$\mathbb{E}_{\lceil Q\rceil_h}\big[h\big(\tfrac{|Y^{(1)}|}{h} - \tfrac{\mathrm{tr}}{h}\big)_+\big] < \tfrac{1}{3}\varepsilon m_Q + 2h < \tfrac{1}{2}\varepsilon m_{\lceil Q\rceil_h}. \tag{3.54}$$

Divide both sides of (3.54) by h, and observe that the continuum word of length $|Y^{(1)}|$ under $\lceil Q\rceil_h$ corresponds to the discrete word of $|Y^{(1)}|/h$ h-letters under Q'_h, to obtain

$$\mathbb{E}_{Q'_h}\big[\big(|Y^{(1)}| - \tfrac{\mathrm{tr}}{h}\big)_+\big] < \tfrac{1}{2}\varepsilon m_{Q'_h}. \tag{3.55}$$

This estimate allows us to use Lemma B.5 in Appendix B, which says that for every $0 < \varepsilon < \frac{1}{2}$,

$$(1-\varepsilon)\Big[H(\Psi_{[Q'_h]_{(\mathrm{tr}/h)}} \mid \nu_h^{\otimes\mathbb{N}}) + b(\varepsilon)\Big] \leq H(\Psi_{Q'_h} \mid \nu_h^{\otimes\mathbb{N}}) \tag{3.56}$$

with $b(\varepsilon) = -2\varepsilon + [\varepsilon\log\varepsilon + (1-\varepsilon)\log(1-\varepsilon)]/(1-\varepsilon)$. However, by (3.51–3.52) we have

$$H(\Psi_{[Q'_h]_{(\mathrm{tr}/h)}} \mid \nu_h^{\otimes\mathbb{N}}) = H(\Psi_{\lceil [Q]_{\mathrm{tr}}\rceil_h,h} \mid \mathscr{W}) = H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}). \tag{3.57}$$

Substitute this relation into (3.56) and use (3.50–3.51), to obtain

$$(1-\varepsilon)\Big[H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}) + b(\varepsilon)\Big] \leq H(\Psi_Q \mid \mathscr{W}). \tag{3.58}$$

Now let $\varepsilon \downarrow 0$ and use that $\lim_{\varepsilon\downarrow 0} b(\varepsilon) = 0$, to obtain (3.45).

3.3.2 Proof of part (2)

Fix $Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with $m_Q = \infty$ and $H(Q \mid Q_{\rho,\mathscr{W}}) < \infty$. We construct $\widetilde{Q}_{\mathrm{tr}} \in \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$, $\mathrm{tr} \in \mathbb{N}$, satisfying (2.25) via a "smoothed truncation" that has the same concatenated word content as its "hard truncation" equivalent. The proof comes in 5 Steps.

Step 1. It will we be convenient to consider the two-sided scenario, i.e., we regard Q as a shift-invariant probability measure on $F^{\mathbb{Z}}$. Define

$$\chi_{\mathrm{tr}} \colon\ F^{\mathbb{Z}}_{0,\mathrm{tr}} \times [0,1]^{\mathbb{Z}} \to F^{\mathbb{Z}}, \qquad \chi_{\mathrm{tr}} \colon\ \big((f_i,\tau_i)_{i\in\mathbb{Z}}, (u_i)_{i\in\mathbb{Z}}\big) \mapsto (\tilde{f}_i, \tilde{\tau}_i)_{i\in\mathbb{Z}}, \tag{3.59}$$

as follows. Put $t_0 = 0$, $t_i = t_{i-1} + \tau_i$, $t_{-i} = t_{-i+1} - \tau_{-i+1}$ for $i \in \mathbb{N}$, and $\varphi = \kappa\big((f_i,\tau_i)_{i\in\mathbb{Z}}\big)$, set

$$\tilde{t}_i = \begin{cases} t_i - u_i & \text{if } \tau_i = \mathrm{tr}, \\ t_i & \text{if } \tau_i < \mathrm{tr}, \end{cases} \tag{3.60}$$

$\tilde{\tau}_i = t_i - t_{i-1}$ and $\tilde{f}_i(\cdot) = \varphi((\,\cdot \wedge \tilde{\tau}_i) + t_{i-1})$ for $i \in \mathbb{Z}$. In words, the total concatenated word content remains unchanged, and if the length of the i-th word τ_i equals tr, then its end-point t_i is moved u_i to the left. Put $\widetilde{Q}_{\mathrm{tr}} = ([Q]_{\mathrm{tr}} \otimes \mathrm{Unif}[0,1]^{\otimes\mathbb{Z}}) \circ \chi_{\mathrm{tr}}^{-1} \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{Z}})$. By construction, $\Psi_{\widetilde{Q}_{\mathrm{tr}}} = \Psi_{[Q]_{\mathrm{tr}}}$ and $m_{\widetilde{Q}_{\mathrm{tr}}} = m_{[Q]_{\mathrm{tr}}}$. In particular,

$$m_{\widetilde{Q}_{\mathrm{tr}}} H(\Psi_{\widetilde{Q}_{\mathrm{tr}}} \mid \mathscr{W}) = m_{[Q]_{\mathrm{tr}}} H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}). \tag{3.61}$$

Step 2. Write $\widetilde{Q}^{\mathrm{ref}}_{\mathrm{tr}} = ([q_{\rho,\mathscr{W}}]^{\otimes\mathbb{Z}}_{\mathrm{tr}} \otimes \mathrm{Unif}[0,1]^{\otimes\mathbb{Z}}) \circ \chi_{\mathrm{tr}}^{-1}$ for the result of the analogous operation on the reference measure $(q_{\rho,\mathscr{W}})^{\otimes\mathbb{Z}}$. We have $\mathrm{w}\text{-}\lim_{\mathrm{tr}\to\infty} \widetilde{Q}_{\mathrm{tr}} = Q$ and $\mathrm{w}\text{-}\lim_{\mathrm{tr}\to\infty}([q_{\rho,\mathscr{W}}]^{\otimes\mathbb{Z}}_{\mathrm{tr}} \otimes \mathrm{Unif}[0,1]^{\otimes\mathbb{Z}}) \circ \chi_{\mathrm{tr}}^{-1} = (q_{\rho,\mathscr{W}})^{\otimes\mathbb{Z}}$, and hence

$$\begin{aligned} \liminf_{\mathrm{tr}\to\infty} H(\widetilde{Q}_{\mathrm{tr}} \mid \widetilde{Q}^{\mathrm{ref}}_{\mathrm{tr}}) &\geq \sup_{\varepsilon>0} \liminf_{\mathrm{tr}\to\infty} \inf_{Q' \in B_\varepsilon(Q)} H(Q' \mid \widetilde{Q}^{\mathrm{ref}}_{\mathrm{tr}}) \\ &\geq H(Q \mid (q_{\rho,\mathscr{W}})^{\otimes\mathbb{Z}}) = \lim_{\mathrm{tr}\to\infty} H([Q]_{\mathrm{tr}} \mid [q_{\rho,\mathscr{W}}]^{\otimes\mathbb{Z}}_{\mathrm{tr}}), \end{aligned} \tag{3.62}$$

where we use Lemma B.2 (2) in the second inequality. (Note: Inspection of the proof of Lemma B.2 (2) shows that the inequality "$\leq$" in (B.10) also holds for Q's that are not product.) The last equality in (3.62) holds because the truncations $[\,\cdot\,]_{\mathrm{tr}}$ form a projective family (see [2, Lemma 8.1]). As specific relative entropy can only decrease under the operation of taking image measures, we have $H(\widetilde{Q}_{\mathrm{tr}} \mid \widetilde{Q}^{\mathrm{ref}}_{\mathrm{tr}}) \leq H([Q]_{\mathrm{tr}} \mid [q_{\rho,\mathscr{W}}]^{\otimes\mathbb{Z}}_{\mathrm{tr}}) \leq H(Q \mid q^{\otimes\mathbb{Z}}_{\rho,\mathscr{W}})$, so $\limsup_{\mathrm{tr}\to\infty} H(\widetilde{Q}_{\mathrm{tr}} \mid \widetilde{Q}^{\mathrm{ref}}_{\mathrm{tr}}) \leq H(Q \mid q^{\otimes\mathbb{Z}}_{\rho,\mathscr{W}})$ and, indeed,

$$\lim_{\mathrm{tr}\to\infty} H(\widetilde{Q}_{\mathrm{tr}} \mid \widetilde{Q}^{\mathrm{ref}}_{\mathrm{tr}}) = \lim_{\mathrm{tr}\to\infty} H([Q]_{\mathrm{tr}} \mid [q_{\rho,\mathscr{W}}]^{\otimes\mathbb{Z}}_{\mathrm{tr}}) = H(Q \mid q^{\otimes\mathbb{Z}}_{\rho,\mathscr{W}}). \tag{3.63}$$

The proof of (2.25) is complete once we show that

$$H(\widetilde{Q}_{\mathrm{tr}} \mid q^{\otimes\mathbb{Z}}_{\rho,\mathscr{W}}) \leq H(\widetilde{Q}_{\mathrm{tr}} \mid \widetilde{Q}^{\mathrm{ref}}_{\mathrm{tr}}) + o(1), \tag{3.64}$$

since, by part (1),

$$\widetilde{I}^{\mathrm{que}}(\widetilde{Q}_{\mathrm{tr}}) = H(\widetilde{Q}_{\mathrm{tr}} \mid q^{\otimes\mathbb{Z}}_{\rho,\mathscr{W}}) + m_{\widetilde{Q}_{\mathrm{tr}}} H(\Psi_{\widetilde{Q}_{\mathrm{tr}}} \mid \mathscr{W}). \tag{3.65}$$

Step 3. It remains to verify (3.64). Note that

$$H(\widetilde{Q}_{\mathrm{tr}} \mid q_{\rho,\mathscr{W}}^{\otimes\mathbb{Z}}) - H(\widetilde{Q}_{\mathrm{tr}} \mid \widetilde{Q}_{\mathrm{tr}}^{\mathrm{ref}}) = \lim_{N\to\infty} \frac{1}{N}\mathbb{E}_{\widetilde{Q}_{\mathrm{tr}}}\left[\log \frac{d\pi_N \widetilde{Q}_{\mathrm{tr}}}{dq_{\rho,\mathscr{W}}^{\otimes N}} - \log \frac{d\pi_N \widetilde{Q}_{\mathrm{tr}}}{d\pi_N \widetilde{Q}_{\mathrm{tr}}^{\mathrm{ref}}}\right]$$

$$= \lim_{N\to\infty} \frac{1}{N}\mathbb{E}_{\widetilde{Q}_{\mathrm{tr}}}\left[\log \frac{d\pi_N \widetilde{Q}_{\mathrm{tr}}^{\mathrm{ref}}}{dq_{\rho,\mathscr{W}}^{\otimes N}}\right], \tag{3.66}$$

and that, by construction, $d\pi_N \widetilde{Q}_{\mathrm{tr}}^{\mathrm{ref}}/dq_{\rho,\mathscr{W}}^{\otimes N}$ is a function of the word lengths $\tilde{\tau}_1, \dots, \tilde{\tau}_N$ only (indeed, because of the i.i.d. property of Brownian increments it easy to see that under both laws the word contents given their lengths are the same, namely, independent pieces of Brownian paths). Write $\widetilde{R}_{\mathrm{tr}}^{\mathrm{ref}}$ for the law of the sequence of word lengths under $\widetilde{Q}_{\mathrm{tr}}^{\mathrm{ref}}$. Then we must show that

$$\limsup_{\mathrm{tr}\to\infty} \lim_{N\to\infty} \frac{1}{N}\mathbb{E}_{\widetilde{Q}_{\mathrm{tr}}}\left[\log \frac{d\pi_N \widetilde{R}_{\mathrm{tr}}^{\mathrm{ref}}}{d\rho^{\otimes N}}(\tilde{\tau}_1, \dots, \tilde{\tau}_N)\right] \le 0. \tag{3.67}$$

Step 4. Denote the density of $\pi_N \widetilde{R}_{\mathrm{tr}}^{\mathrm{ref}}$ with respect to Lebesgue measure on $\mathbb{R}_+^N$ by $\widetilde{f}_{\mathrm{tr},N}^{\mathrm{ref}}$. Consider fixed $\tilde{\tau}_1, \dots, \tilde{\tau}_N$, and decompose into maximal stretches of $\tilde{\tau}_i$'s with values in $(\mathrm{tr}-1, \mathrm{tr}+1)$ (note that under χ_{tr} no word can become longer than $\mathrm{tr}+1$, while when $\tilde{\tau}_i < \mathrm{tr}-1$ the corresponding word is not truncated, i.e., $\tilde{t}_i = t_i$). Thus, there are $0 \le M < N$, $i_1' \le j_2' < i_2' \le j_2' < \dots < i_M' \le j_M' \le N$ such that $\{1 \le i \le N\colon\ \tilde{\tau}_i \in (\mathrm{tr}-1, \mathrm{tr}+1)\} = \cup_{k=1}^M [i_k', j_k'] \cap \mathbb{N}$. Observe that, by construction, $\widetilde{f}_{\mathrm{tr},N}^{\mathrm{ref}}(\tilde{\tau}_1, \dots, \tilde{\tau}_N)$ can be decomposed into a product of $\prod_{j\colon \tilde{\tau}_j \le \mathrm{tr}-1} \bar{\rho}(\tilde{\tau}_j)$ and M further factors involving the $\tilde{\tau}_i$'s from these stretches, where the k-th factor depends only on $(\tilde{\tau}_i\colon\ i_k' \le i \le j_k')$. We claim that

$$\frac{\widetilde{f}_{\mathrm{tr},N}^{\mathrm{ref}}(\tilde{\tau}_1, \dots, \tilde{\tau}_N)}{\prod_{j=1}^N \bar{\rho}(\tilde{\tau}_j)} \le \prod_{k=1}^M \left(C_1 \mathrm{tr}^{1+\epsilon}\right)^{j_k' - i_k' + 1} = \left(C_1 \mathrm{tr}^{1+\epsilon}\right)^{\#\{1 \le i \le N\colon\ \tilde{\tau}_i > \mathrm{tr}-1\}} \tag{3.68}$$

for some $C_1 = C_1(\rho) < \infty$ and $\epsilon = \epsilon(\rho) \in [0,1]$ uniformly in tr for tr sufficiently large. To see why (3.68) holds, consider for example the first stretch and assume for simplicity that $i_1' = 1 < j_1'$ and that we know that the 0-th word is not truncated (i.e., $\tilde{t}_0 = t_0 = 0$). Let $\ell \le j_1' + 1$, and pretend we know that the first $\ell - 1$ words are truncated (i.e., $\tau_1 = \dots = \tau_{\ell-1} = \mathrm{tr}$), while the ℓ-th word is not ($\tau_\ell < \mathrm{tr}$). Then $\tilde{\tau}_1 = \mathrm{tr} - u_1$ and $\tilde{\tau}_i = \mathrm{tr} - u_i + u_{i-1}$ for $2 \le i \le \ell - 1$, and so $u_i = \sum_{j=1}^i (\mathrm{tr} - \tilde{\tau}_j)$ for $1 \le i \le \ell - 1$ and $\tau_\ell = \tilde{\tau}_\ell - u_{\ell-1} = \tilde{\tau}_\ell - \sum_{j=1}^{\ell-1}(\mathrm{tr} - \tilde{\tau}_j)$. This case contributes to $\widetilde{f}_{\mathrm{tr},\ell}^{\mathrm{ref}}$ the term

$$\rho([\mathrm{tr}, \infty))^{\ell-1} \bar{\rho}\Big(\tilde{\tau}_\ell - \textstyle\sum_{j=1}^{\ell-1}(\mathrm{tr} - \tilde{\tau}_j)\Big) \displaystyle\prod_{i=1}^{\ell-1} 1_{[0,1]}\Big(\textstyle\sum_{j=1}^i (\mathrm{tr} - \tilde{\tau}_j)\Big). \tag{3.69}$$

Note that, by (1.1), we have $(3.69)/\prod_{j=1}^\ell \bar{\rho}(\tilde{\tau}_j) \le C_2 (C_3 \mathrm{tr}^{1+\epsilon})^{\ell-1}$ for some $C_2 = C_2(\rho), C_3 = C_3(\rho) < \infty$ and $\epsilon = \epsilon(\rho) \in [0,1]$ uniformly in tr for tr sufficiently large. The contribution of any given stretch of length $j_k' - i_k' + 1$ can be written as a sum of at most $2^{j_k' - i_k' + 1}$ cases where the indices of the truncated words are specified. Each such case can be estimated by a suitable product of terms as in (3.69). Furthermore, outside the stretches the words are necessarily untruncated and thus contribute $\bar{\rho}(\tilde{\tau}_i)$ to $\widetilde{f}_{\mathrm{tr},N}^{\mathrm{ref}}$, which cancels with the corresponding term in $\rho^{\otimes N}$.

Step 5. From (3.68) and the shift-invariance of $\widetilde{Q}_{\rm tr}$ we obtain that

$$\lim_{N\to\infty}\frac{1}{N}\mathbb{E}_{\widetilde{Q}_{\rm tr}}\left[\log\frac{d\pi_N\widetilde{R}^{\rm ref}_{\rm tr}}{d\rho^{\otimes N}}(\tilde\tau_1,\dots,\tilde\tau_N)\right]\le C(1+\log{\rm tr})Q(\tau_1>{\rm tr}-1). \tag{3.70}$$

Now, $h(\mathscr{L}_Q(\tau_1)\mid\rho)\le H(Q\mid q^{\mathbb{N}}_{\rho,\mathscr{W}})<\infty$ by assumption. Because of (1.1), this implies that $\mathbb{E}_Q[\log(\tau_1)]<\infty$, and hence that $Q(\tau_1>{\rm tr})=o(1/\log{\rm tr})$. Therefore (3.70) implies (3.67).
□

4 Removal of Assumptions (2.14)–(2.15)

We give a brief sketch of the proof only, leaving the details to the reader. Assumptions (2.14)–(2.15) are satisfied when $\bar\rho$ satisfies (1.2) and varies regularly at ∞ with index α. The latter condition is stronger than (1.1). To prove the claim under (1.1) alone, note that for every $\delta>0$ and $\alpha'<\alpha$ there exists a probability density $\bar\rho'=\bar\rho'(\delta,\alpha')$ such that $\bar\rho\le(1+\delta)\bar\rho'$, $\bar\rho'$ varies regularly at ∞ with index α', and $\bar\rho'(t)dt$ converges weakly to $\bar\rho(t)dt$ as $\delta\downarrow 0$ and $\alpha'\uparrow\alpha$. Since the quenched LDP holds for $\bar\rho'$, we can proceed similarly as in [2, Sections 3.6 and 5] to get the quenched LDP for $\bar\rho$.

More precisely, for $B\subset\mathcal{P}^{\rm inv}(F^{\mathbb{N}})$ we may write (recall (1.6) and (2.31))

$$P(R_N\in B\mid X)=\int_{0\le t_1<\cdots<t_N<\infty}dt_1\cdots dt_N\,\bar\rho(t_1)\,\bar\rho(t_2-t_1)\cdots\bar\rho(t_N-t_{N-1})\times 1_B\big(R_{N;t_1,\dots,t_N}(X)\big), \tag{4.1}$$

and estimate $\bar\rho(t_1)\le(1+\delta)\bar\rho'(t_1)$, etc., to get $P(R_N\in B\mid X)\le(1+\delta)^N\,P'(R_N\in B\mid X)$, where P,P' have $\bar\rho,\bar\rho'$ as excursion length distributions. Let $\mathcal{C}\subset\mathcal{P}^{\rm inv}(F^{\mathbb{N}})$ be a closed set, and let $\mathcal{C}^{(\varepsilon)}$ be its ε-blow-up. Then the LDP upper bound for $\bar\rho'$ gives

$$\limsup_{N\to\infty}\frac{1}{N}\log P(R_N\in\mathcal{C}^{(\varepsilon)}\mid X)\le\log(1+\delta)-\inf_{Q\in\mathcal{C}^{(\varepsilon)}}I^{\rm que}_{\bar\rho'}(Q)\qquad X\text{-a.s.}, \tag{4.2}$$

where the lower index $\bar\rho'$ indicates the excursion length distribution. Let $\delta\downarrow 0$ and $\alpha'\uparrow\alpha$, and use Lemma B.2 (2), to get

$$\limsup_{N\to\infty}\frac{1}{N}\log P(R_N\in\mathcal{C}^{(\varepsilon)}\mid X)\le-\inf_{Q\in\mathcal{C}^{(2\varepsilon)}}I^{\rm que}_{\bar\rho}(Q)\qquad X\text{-a.s.} \tag{4.3}$$

Finally, let $\varepsilon\downarrow 0$ and use the lower semi-continuity of $I^{\rm que}_{\bar\rho}$ to get the LDP upper bound for $\bar\rho$.

An analogous argument works for the LDP lower bound: Now we pick $\alpha'>\alpha$, $\delta>0$ and a probability density $\bar\rho'=\bar\rho'(\delta,\alpha')$ such that $\bar\rho\ge(1-\delta)\bar\rho'$, and $\bar\rho'$ satisfies the same conditions as above. Arguing as before, we obtain for any open $\mathcal{O}\subset\mathcal{P}^{\rm inv}(F^{\mathbb{N}})$,

$$\liminf_{N\to\infty}\frac{1}{N}\log P(R_N\in\mathcal{C}^{(\varepsilon)}\mid X)\ge-\inf_{Q\in\mathcal{O}}I^{\rm que}_{\bar\rho}(Q)\qquad X\text{-a.s.} \tag{4.4}$$

5 Proof of Theorems 1.3–1.4

We again give a brief sketch of the proofs only, leaving many details to the reader.

Theorem 1.3(a), which says that for $\alpha = 1$ the quenched rate function coincides with the annealed rate function, can be proved as follows: Since the claimed LDP upper bound holds automatically by the annealed LDP, it suffices to verify the matching lower bound. For this we can argue as in the proof of the lower bound in Section 4. For any $\alpha' > 1$ and $\delta > 0$ we can approximate $\bar{\rho}$ by a suitable $\bar{\rho}' = \bar{\rho}'(\delta, \alpha')$ such that $\bar{\rho} \geq (1-\delta)\bar{\rho}'$. Then, using Theorem 1.2 with $\bar{\rho}'$ and taking $\delta \downarrow 0$, $\alpha' \downarrow 1$, we see that for any open $\mathcal{O} \subset \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$,

$$\liminf_{N\to\infty} \frac{1}{N} \log P(R_N \in \mathcal{C}^{(\varepsilon)} \mid X) \geq - \inf_{Q \in \mathcal{O} \cap \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})} I^{\mathrm{ann}}(Q) \qquad X\text{-a.s.} \tag{5.1}$$

(recall (1.15)). Finally note that any $Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with $H(Q \mid Q_{\rho,\mathscr{W}}) < \infty$ can be approximated by a sequence $(Q_n) \subset \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$ in such a way that $H(Q_n \mid Q_{\rho,\mathscr{W}}) \to H(Q \mid Q_{\rho,\mathscr{W}})$ to obtain the claim (using for example a "smoothed truncation" operation similar to Section 3.3.2).

Theorem 1.3(b), which says that for $\alpha = \infty$ the quenched rate function coincides with the annealed rate function on the set $\{Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})\colon \lim_{\mathrm{tr}\to\infty} m_{[Q]_{\mathrm{tr}}} H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}) = 0\}$ and is infinite elsewhere, follows from arguments analogous to [2, Section 7, Part (b)]: For the upper bound, we can pick arbitrarily large $\alpha' > 1$ and approximate $\bar{\rho} \leq (1+\delta)\bar{\rho}'$ with the help of a suitable probability density $\bar{\rho}'$ which has decay exponent α'. Using Theorem 1.2 with $\bar{\rho}'$ and taking $\alpha' \uparrow \infty$, $\delta \downarrow 0$, we see that the upper bound holds with the claimed form of the rate function. For the matching lower bound we can trace through the proof of the lower bound contained in Theorem 1.2 but replacing our "coarse-graining work horses" Proposition 2.1 and Corollary 2.2 (which rely on [2, Cor. 1.6]) by versions that are suitable for $\alpha = \infty$ (which rely on [2, Thm. 1.4 (b)] instead), still using a suitable truncation approximation of the quenched rate function analogous to the one proven in Proposition 2.5. This constitutes a way of rigorously implementing the "first long string strategy" from [2, Section 4], as explained in the heuristic given in item 0 of Section 1.3, through the coarse-graining approximation.

Theorem 1.4 follows from Theorem 1.3(b) via an observation that is the analogue of [1, Lemma 6]: subject to the exponential tail condition in (1.19), any $Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$ with $H(Q \mid Q_{\rho,\mathscr{W}}) < \infty$ necessarily has $m_Q < \infty$. Because of this observation we can argue as follows. If $m_Q < \infty$, then $\lim_{\mathrm{tr}\to\infty} m_{[Q]_{\mathrm{tr}}} = m_Q$ and $\lim_{\mathrm{tr}\to\infty} \Psi_{[Q]_{\mathrm{tr}}} = \Psi_Q$ by dominated convergence (recall (1.10)), which in turn imply that $\liminf_{\mathrm{tr}\to\infty} m_{[Q]_{\mathrm{tr}}} H(\Psi_{[Q]_{\mathrm{tr}}} \mid \mathscr{W}) = m_Q H(\Psi_Q \mid \mathscr{W})$, as shown in Lemma B.5 in Appendix B. The limit is zero if and only if $\Psi_Q = \mathscr{W}$, which by (1.22) holds if and only if $Q \in \mathcal{R}_{\mathscr{W}}$. This explains the link between (1.18) and (1.20).

A Basic facts about metrics on path space

We metrise F, defined in (1.4) (and $F_h \subset F$ defined in (2.3)) as follows. Let $d_S(\phi_1, \phi_2)$ be a metric on $C([0,\infty))$ that generates Skorohod's J_1-topology on $D([0,\infty)) \supset C([0,\infty))$, allowing for a certain amount of "rubber time" (see e.g. Ethier and Kurtz [9, Section 3.5 and Eqs. (5.1–5.3)])

$$d_S(\phi_1, \phi_2) = \inf_{\lambda \in \Lambda} \left\{ \gamma(\lambda) \vee \int_0^\infty e^{-u} \sup_{t \geq 0} \big|\phi_1(t \wedge u) - \phi_2(\lambda(t) \wedge u)\big| \, du \right\}, \tag{A.1}$$

where Λ is the set of Lipschitz-continuous bijections from $[0,\infty)$ into itself and

$$\gamma(\lambda) = \sup_{0 \leq s < t < \infty} \Big| \log \frac{\lambda(t) - \lambda(s)}{t-s} \Big|. \tag{A.2}$$

With

$$d_F(y_1, y_2) = |t_1 - t_2| + d_S(\phi_1, \phi_2) \tag{A.3}$$

for $y_i = (t_i, \phi_i) \in F$, (F, d_F) becomes complete and separable, and the same holds for (F_h, d_F) for any $h > 0$.

Remark. We might at first be inclined to metrise F in a more straightforward way than (A.3), e.g. via

$$d_F^{\text{first}}(y_1, y_2) = |t_1 - t_2| + \|\phi_1 - \phi_2\|_\infty, \quad y_i = (t_i, \phi_i) \in F,\ i = 1, 2. \tag{A.4}$$

However, if we would use Lipschitz functions with d_F replaced by d_F^{first} in (2.16), then in the analogue of Lemma 2.7 we would be forced to use terms of the form $\sup_{s\geq 0} |\varphi(s + t \wedge t') - \varphi(s + ih \wedge jh)|$ in the right-hand side. When used for $\varphi = X$ (a realisation of Brownian motion as in Proposition 2.3), this would in turn force us to control the increments of the Brownian motion not only locally near the beginning and the end of each loop, but uniformly inside loops. In fact, it seems plausible that an analogue of Proposition 2.3 where d_F is replaced by d_F^{first} actually fails. Furthermore, note that we cannot arrange d_S in such a way that, for $\phi \in C([0, \infty))$, $h > 0$, $t_1 \leq t_1' < t_2 \leq t_2'$ with $|t_1' - t_1| \leq h$, $|t_2' - t_2| \leq h$,

$$d_S\big(\phi((t_1 + \cdot) \wedge t_2), \phi((t_1' + \cdot) \wedge t_2')\big) \leq 2h + \sup_{t_1 \leq s \leq t_1'} |\phi(s) - \phi(t_1')| + \sup_{t_2 \leq s \leq t_2'} |\phi(s) - \phi(t_2')|. \tag{A.5}$$

This is why in Lemma 2.7 we need the freedom to use an extra k and to "look in a neighbourhood of the cut-points of size kh".

B Basic facts about specific relative entropy

In Section B.1 we recall the definition of (specific) relative entropy of two probability measures. In Section B.2 we prove various approximation results for (specific) relative entropy, which were used heavily in Sections 3. Especially the parts with Ψ_Q require care because of the delicate nature of the word concatenation map $Q \mapsto \Psi_Q$. The latter is looked at in closer detail in Section B.3.

B.1 Definitions

For μ, ν probability measures on a measurable space $(S, \mathscr{S})$,

$$h(\mu \mid \nu) = \begin{cases} \int_S (\log \frac{d\mu}{d\nu})\, d\mu, & \text{if } \mu \ll \nu, \\ \infty, & \text{otherwise,} \end{cases} \tag{B.1}$$

is the relative entropy of μ w.r.t. ν. When the measurable space is a Polish space E equipped with its Borel-σ-algebra, we also have the representation (see e.g. [7, Lemma 6.2.13])

$$h(\mu \mid \nu) = \sup_{f \in C_b(E)} \Big\{ \int f\, d\mu - \log \int e^f\, d\nu \Big\} = \sup_{\substack{f\colon E \to \mathbb{R} \\ \text{bounded measurable}}} \Big\{ \int f\, d\mu - \log \int e^f\, d\nu \Big\} \tag{B.2}$$

(and if $\mu \ll \nu$ with a bounded and uniformly positive density, then the supremum in the right-hand side is achieved by $f = \log d\mu/d\nu$).

Equation (B.2) implies the entropy inequality

$$\mu(A) \le \frac{\log 2 + h(\mu \mid \nu)}{\log[1 + 1/\nu(A)]} \tag{B.3}$$

by choosing $f = \alpha 1_A$ and $\alpha = \log[1 + 1/\nu(A)]$ (see e.g. Kipnis and Landim [13, Appendix 1, Proposition 8.2]).

For $Q \in \mathcal{P}^{\mathrm{inv}}(F^{\mathbb{N}})$,

$$H(Q \mid (q_{\rho,\mathscr{W}})^{\otimes \mathbb{N}}) = \lim_{N\to\infty} \frac{1}{N} h\big(\pi_N Q \mid (q_{\rho,\mathscr{W}})^{\otimes N}\big) = \sup_{N\in\mathbb{N}} \frac{1}{N} h\big(\pi_N Q \mid (q_{\rho,\mathscr{W}})^{\otimes N}\big) \tag{B.4}$$

with π_N the projection onto the first N words, is the specific relative entropy of Q w.r.t. $(q_{\rho,\mathscr{W}})^{\otimes \mathbb{N}}$. Similarly, using the canonical filtration $(\mathscr{F}_t^C)_{t\ge 0}$ on $C([0,\infty))$, for a probability measure Ψ on $C([0,\infty))$ with stationary increments we denote by

$$H(\Psi \mid \mathscr{W}) = \lim_{t\to\infty} \frac{1}{t} h\big(\Psi_{|\mathscr{F}_t^C} \mid \mathscr{W}_{|\mathscr{F}_t^C}\big) = \sup_{t>0} \frac{1}{t} h\big(\Psi_{|\mathscr{F}_t^C} \mid \mathscr{W}_{|\mathscr{F}_t^C}\big) \tag{B.5}$$

the specific relative entropy w.r.t. Wiener measure. See Appendix C for a proof of (B.5).

B.2 Approximations

Let E be a Polish space. Equip $\mathcal{P}(E)$ with the weak topology (suitably metrised). $E^{\mathbb{N}}$ carries the product topology, and the set of shift-invariant probability measures $\mathcal{P}^{\mathrm{inv}}(E^{\mathbb{N}})$ carries the weak topology (also suitably metrised).

B.2.1 Blocks

For $M \in \mathbb{N}$ and $r \in \mathcal{P}(E^M)$, denote by $r^{\otimes \mathbb{N}} \in \mathcal{P}(E^{\mathbb{N}})$ the law of an infinite sequence obtained by concatenating M-blocks drawn independently from r (i.e., we identify $(E^M)^{\mathbb{N}}$ and $E^{\mathbb{N}}$ in the obvious way), and write

$$\mathsf{sblock}_M(r) = \frac{1}{M} \sum_{j=0}^{M-1} r^{\otimes \mathbb{N}} \circ (\theta^j)^{-1} \;\in \mathcal{P}^{\mathrm{inv}}(E^{\mathbb{N}}) \tag{B.6}$$

for its stationary mean.

Lemma B.1. *For $Q = q^{\otimes \mathbb{N}} \in \mathcal{P}^{\mathrm{inv}}(E)$ and $r \in \mathcal{P}(E^M)$,*

$$H\big(\mathsf{sblock}_M(r) \mid Q\big) = \frac{1}{M} h\big(r \mid \pi_M Q\big). \tag{B.7}$$

Moreover, for any $R \in \mathcal{P}^{\mathrm{inv}}(E)$,

$$\operatorname*{w-lim}_{M\to\infty} \mathsf{sblock}_M\big(\pi_M R\big) = R. \tag{B.8}$$

Proof. This proof is standard. Equation (B.7) follows from the results in Gray [12, Section 8.4, see Theorem 8.4.1] by observing that $\mathsf{sblock}_M(r)$ is the asymptotically mean stationary measure of $r^{\otimes \mathbb{N}}$. It is also contained in Föllmer[11, Lemma 4.8], or can be proved with "bare hands" by explicitly spelling out $d\pi_N \mathsf{sblock}_M(r)/dq^{\otimes N}$ for $N \gg M$ and using suitable concentration arguments under $q^{\otimes N}$ as $N \to \infty$. Equation (B.8) is obvious from the definition of weak convergence. □

B.2.2 Change of reference measure

Lemma B.2. (1) *Let* $\nu, \nu_1, \nu_2, \ldots \in \mathcal{P}(E)$ *with* $\text{w-lim}_{n\to\infty} \nu_n = \nu$. *Then*

$$h(\mu \mid \nu) = \lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} \inf_{\mu' \in B_\varepsilon(\mu)} h(\mu' \mid \nu_n), \quad \mu \in \mathcal{P}(E). \tag{B.9}$$

(2) *Let* $Q = q^{\otimes\mathbb{N}}, Q_1 = q_1^{\otimes\mathbb{N}}, Q_2 = q_2^{\otimes\mathbb{N}}, \ldots \in \mathcal{P}^{\text{inv}}(E^{\mathbb{N}})$ *be product measures with* $\text{w-lim}_{n\to\infty} Q_n = Q$. *Then*

$$H(R \mid Q) = \lim_{\varepsilon \downarrow 0} \limsup_{n\to\infty} \inf_{R' \in B_\varepsilon(R)} H(R' \mid Q_n), \quad R \in \mathcal{P}^{\text{inv}}(E^{\mathbb{N}}). \tag{B.10}$$

Proof. (1) Denote the term in the right-hand side of (B.9) by $\tilde{h}(\mu)$. Let $f \in C_b(E)$, $\delta > 0$. We can find $\varepsilon_0 > 0$ and $n_0 \in \mathbb{N}$ such that

$$\forall\, 0 < \varepsilon \le \varepsilon_0,\ \mu' \in B_\varepsilon(\mu)\colon \quad \int_E f\, d\mu' \ge \int_E f\, d\mu - \frac{\delta}{2}, \tag{B.11}$$

$$\forall\, n \ge n_0\colon \quad \log \int_E e^f\, d\nu_n \le \log \int_E e^f\, d\nu + \frac{\delta}{2}. \tag{B.12}$$

Therefore, for $0 < \varepsilon \le \varepsilon_0$ and $n \ge n_0$,

$$\inf_{\mu' \in B_\varepsilon(\mu)} h(\mu' \mid \nu_n) \ge \int_E f\, d\mu' - \log \int_E e^f\, d\nu_n \ge \int_E f\, d\mu - \log \int_E e^f\, d\nu - \delta. \tag{B.13}$$

Now optimise over f and take $\delta \downarrow 0$, to obtain $\tilde{h}(\mu) \ge h(\mu \mid \nu)$ via (B.2).

For the reverse inequality, we may without loss of generality assume that $h(\mu \mid \nu) = \int_E \varphi \log \varphi\, d\nu < \infty$, where $\varphi = d\mu/d\nu \ge 0$ is in $L^1(\nu)$. Then for any $\delta > 0$ we can find a $\widetilde{\varphi} \ge 0$ in $C_b(E) \cap L^1(\nu)$ such that $\int_E \widetilde{\varphi}\, d\nu = 1$ and

$$\int_E \big|\widetilde{\varphi} - \varphi\big|\, d\nu < \delta, \quad \int_E \big|\widetilde{\varphi} \log \widetilde{\varphi} - \varphi \log \varphi\big|\, d\nu < \delta. \tag{B.14}$$

Note that $\lim_{n\to\infty} \int_E \widetilde{\varphi}\, d\nu_n = 1$, and let $\widetilde{\varphi}_n = \widetilde{\varphi} / \int \widetilde{\varphi}\, d\nu_n$ and $\mu_n = \widetilde{\varphi}_n \nu_n$. Then, for $g \in C_b(E)$,

$$\begin{aligned} \Big| \int_E g\, d\mu_n - \int_E g\, d\mu \Big| &= \Big| \frac{1}{\int_E \widetilde{\varphi}\, d\nu_n} \int_E g\widetilde{\varphi}\, d\nu_n - \int_E g\varphi\, d\nu \Big| \\ &\le \Big| \frac{1}{\int_E \widetilde{\varphi}\, d\nu_n} - 1 \Big|\, \|g\widetilde{\varphi}\|_\infty + \Big| \int_E g\widetilde{\varphi}\, d\nu_n - \int_E g\widetilde{\varphi}\, d\nu \Big| + \Big| \int_E g(\widetilde{\varphi} - \varphi)\, d\nu \Big|, \end{aligned} \tag{B.15}$$

which can be made arbitrarily small by choosing δ small enough and n large enough. In particular, for any $\varepsilon > 0$ we can choose δ, $\widetilde{\varphi}$ and n_0 such that $\mu_n \in B_\varepsilon(\mu)$ for $n \ge n_0$. Hence

$$\begin{aligned} \limsup_{n\to\infty} \inf_{\mu' \in B_\varepsilon(\mu)} h(\mu' \mid \nu_n) &\le \limsup_{n\to\infty} h(\mu_n \mid \nu_n) \\ &= \limsup_{n\to\infty} \int_E \widetilde{\varphi}_n \log \widetilde{\varphi}_n\, d\nu_n = \int_E \widetilde{\varphi} \log \widetilde{\varphi}\, d\nu \le h(\mu \mid \nu) + \delta, \end{aligned} \tag{B.16}$$

and letting $\delta \downarrow 0$ we $\tilde{h}(\mu) \le h(\mu \mid \nu)$.

(2) Recall that for $R \in \mathcal{P}^{\text{inv}}(E^{\mathbb{N}})$ and Q a product measure,

$$\lim_{N\to\infty} \frac{1}{N} h\left(\pi_N R \mid \pi_N Q\right) = H(R \mid Q) = \sup_{N \in \mathbb{N}} \frac{1}{N} h\left(\pi_N R \mid \pi_N Q\right). \tag{B.17}$$

Denote the expression in the right-hand side of (B.10) by $\tilde{H}(R)$. Fix $N \in \mathbb{N}$. Since for each $\varepsilon > 0$, we have $B_{\varepsilon'}(R) \subset \{R' \colon \pi_N R' \in B_\varepsilon(\pi_N R)\}$ for ε' sufficiently small we also have

$$\lim_{\varepsilon' \downarrow 0} \limsup_{n\to\infty} \inf_{R' \in B_{\varepsilon'}(R)} H(R' \mid Q_n) \geq \limsup_{n\to\infty} \inf_{\mu' \in B_\varepsilon(\pi_N R)} \frac{1}{N} h(\mu' \mid \pi_N Q_n). \tag{B.18}$$

Let $\varepsilon \downarrow 0$ and use Part (1), to see that $\tilde{H}(R) \geq \frac{1}{N} h(\pi_N R \mid \pi_N Q)$ for any N. Hence also $\tilde{H}(R) \geq H(R \mid Q)$.

For the reverse inequality, we may w.l.o.g. assume that $H(R \mid Q) < \infty$. Fix $\varepsilon > 0$ and $\delta > 0$. There is an $N \in \mathbb{N}$ such that $H(R \mid Q) \leq \frac{1}{N} h\big(\pi_N R \mid \pi_N Q\big) + \delta$, and since $\pi_N R \ll \pi_N Q = q^{\otimes N}$ we can find a continuous, bounded and uniformly positive function $f_N \colon E^N \to [0,\infty)$ such that $\int_E f_N \, dq^{\otimes N} = 1$, $\int_E f_N \log f_N \, dq^{\otimes N} \leq h\big(\pi_N R \mid \pi_N Q\big) + N\delta$ and $\tilde{R}_N \in B_{\varepsilon/2}(R)$, where $\tilde{R}_N = \mathsf{sblock}_N\big(f_N \, q^{\otimes N}\big) \in \mathcal{P}^{\mathrm{inv}}(E^{\mathbb{N}})$ (see Lemma B.1). By (B.7), we have

$$H(\tilde{R}_N \mid Q) = \frac{1}{N} \int_E f_N \log f_N \, dq^{\otimes N}. \tag{B.19}$$

Now put $f_{N,n} = f_N / \int_E f_N \, q_n^{\otimes N}$, and define $\tilde{R}_{N,n} = \mathsf{sblock}_N\big(f_{N,n} \, q^{\otimes N}\big)$ as the "stationary version" of $(f_{N,n} \, q_n^{\otimes N})^{\otimes \mathbb{N}}$. In particular, $H(\tilde{R}_{N,n} \mid Q_n) = \frac{1}{N} \int f_{N,n} \log f_{N,n} \, dq_n^{\otimes N}$. Since f_N is continuous, we have $\tilde{R}_{N,n} \in B_\varepsilon(R)$ and $\int_E f_{N,n} \log f_{N,n} \, dq_n^{\otimes N} \leq H(R \mid Q) + 3\delta$ for n large enough. Hence

$$\limsup_{n\to\infty} \inf_{R' \in B_\varepsilon(R)} H(R' \mid Q_n) \leq \limsup_{n\to\infty} H(\tilde{R}_{N,n} \mid Q_n) \leq H(R \mid Q) + 4\delta. \tag{B.20}$$

Now let $\delta \downarrow 0$ followed by $\lim_{\varepsilon \downarrow 0}$ to conclude the proof. □

B.2.3 Existence of sharp coarse-graining approximations to the quenched rate function

The following lemma was used in the proof of Lemma 3.3.

Lemma B.3. *Let $Q \in \mathcal{P}^{\mathrm{fin}}(F^{\mathbb{N}})$ with $H(Q \mid Q_{\rho,\mathscr{W}}) < \infty$. There exist a sequence $(h_n)_{n\in\mathbb{N}}$ with $h_n > 0$ and $\lim_{n\to\infty} h_n = 0$ and a sequence $(Q_n')_{n\in\mathbb{N}}$ with $Q_n' \in \mathcal{P}^{\mathrm{fin}}(\widetilde{E}_{h_n}^{\mathbb{N}})$ and $\text{w-}\lim_{n\to\infty} Q_n' = Q$ such that $\limsup_{n\to\infty} I_{h_n}^{\mathrm{que}}(Q_n') \leq I^{\mathrm{que}}(Q)$. The same holds with F replaced by $F_{0,\mathrm{tr}}$ and $\widetilde{E}_{h_n}$ replaced by $\widetilde{E}_{h_n,\mathrm{tr}}$.*

Proof. Recall the definition of $\lceil Q \rceil_h$ in Step 2 of the proof of part (1) of Proposition 2.5 (see page 21). For any $N \in \mathbb{N}$, we have

$$h(\pi_N \lceil Q \rceil_h \mid \pi_N \lceil Q_{\rho,\mathscr{W}} \rceil_h) \leq h(\pi_N Q \mid \pi_N Q_{\rho,\mathscr{W}}) \leq N\, H(Q \mid Q_{\rho,\mathscr{W}}). \tag{B.21}$$

The second inequality follows from (B.4). For the first inequality, use the fact that the construction of $\lceil Q \rceil_h$ can be implemented as a deterministic function of the pair of random variables (Y, U), together with the fact that relative entropy can only decrease when image measures are taken. Write $\hat{\tau}_i = (\tilde{T}_i - \tilde{T}_{i-1})/h$, $i \in \mathbb{N}$. Since letters both under $\lceil Q_{\rho,\mathscr{W}} \rceil_h$ and under $Q_{\lceil \rho \rceil_h,\mathscr{W}}$ are constructed from a Brownian path that is independent of the word lengths, we have (recall 2.8)

$$1(\hat{\tau}_1 = k_1, \ldots, \hat{\tau}_N = k_N) \, \frac{d\pi_N \lceil Q_{\rho,\mathscr{W}} \rceil_h}{d\pi_N Q_{\lceil \rho \rceil_h,\mathscr{W}}} = \frac{(\pi_N \lceil Q_{\rho,\mathscr{W}} \rceil_h)\big(\hat{\tau}_1 = k_1, \ldots, \hat{\tau}_N = k_N\big)}{\prod_{\ell=1}^N \lceil \rho \rceil_h (h k_\ell)} \tag{B.22}$$

with

$$
\begin{aligned}
&(\pi_N \lceil Q_{\rho,\mathscr{W}} \rceil_h)\big(\hat\tau_1 = k_1, \dots, \hat\tau_N = k_N\big) \\
&= \int_{[0,1]} du \int_0^\infty \bar\rho(t_1) dt_1 \int_{t_1}^\infty \bar\rho(t_2 - t_1) d(t_2 - t_1) \cdots \int_{t_{N-1}}^\infty \bar\rho((t_N - t_{N-1})) d(t_N - t_{N-1}) \\
&\qquad\qquad \times \prod_{\ell=1}^N 1_{(h(\bar k_\ell - 1 + u), h(\bar k_\ell + u)]}(t_\ell),
\end{aligned}
\tag{B.23}
$$

where $\bar k_\ell = k_1 + \cdots + k_\ell$. Thus, by (2.13–2.15),

$$
\sup_{N\in\mathbb{N}} \frac{1}{N} E_{\lceil Q \rceil_h} \left[\Big| \log \frac{d\pi_N \lceil Q_{\rho,\mathscr{W}} \rceil_h}{d\pi_N Q_{\lceil \rho \rceil_h, \mathscr{W}}} \Big| \right] \leq r_Q(h) \tag{B.24}
$$

with

$$
r_Q(h) = \eta_n \lceil Q \rceil_h (\hat\tau_1 \in \bar A_n) + \eta_0 \lceil Q \rceil_h (\hat\tau_1 \notin \bar A_n), \qquad h = 2^{-n}, \tag{B.25}
$$

where $\bar A_n \subset (s_*, \infty)$ is the set obtained from A_n by removing pieces of length 2^{-n} from its edges (i.e., $\bar A_n$ is the 2^{-n}-interior of A_n). But $\lim_{n\to\infty} \lceil Q \rceil_{2^{-n}} (\hat\tau_1 \notin \bar A_n) = 0$ because A_n fills up (s_*, ∞) as $n \to \infty$. Since $\lim_{n\to\infty} \eta_n = 0$, we get $\lim_{h\downarrow 0} r_Q(h) = 0$. Combining (B.21–B.24), we obtain that

$$
H(\lceil Q \rceil_h \mid Q_{\lceil \rho \rceil_h, \mathscr{W}}) = \sup_{N\in\mathbb{N}} \frac{1}{N} h(\pi_N \lceil Q \rceil_h \mid \pi_N Q_{\lceil \rho \rceil_h, \mathscr{W}}) \leq H(Q \mid Q_{\rho,\mathscr{W}}) + r_Q(h) \tag{B.26}
$$

and, finally,

$$
\begin{aligned}
&\limsup_{h\downarrow 0} H(\lceil Q \rceil_h \mid Q_{\lceil \rho \rceil_h, \mathscr{W}}) + (\alpha - 1) m_{\lceil Q \rceil_h} H(\Psi_{\lceil Q \rceil_h, h} \mid \mathscr{W}) \\
&\qquad \leq H(Q \mid Q_{\rho,\mathscr{W}}) + (\alpha - 1) m_Q H(\Psi_Q \mid \mathscr{W}).
\end{aligned}
\tag{B.27}
$$

The truncated case, where F is replaced by $F_{0,\mathrm{tr}}$, etc., can be handled analogously. □

B.2.4 Approximation of Ψ_Q

The approximation in (1.11) is stronger than just weak convergence.

Lemma B.4. *For $Q \in \mathcal{P}^{\mathrm{inv,fin}}(F^{\mathbb{N}})$,*

$$
\lim_{T\to\infty} \sup_{A \subset C[0,\infty) \text{ measurable}} \left| \Psi_Q(A) - \frac{1}{T} \int_0^T \big(\kappa(Q) \circ (\theta^s)^{-1}\big)(A)\, ds \right| = 0, \tag{B.28}
$$

i.e., the convergence in (1.11) *holds in total variation.*

Proof. Note that, by shift-invariance,

$$
\Psi_Q(A) = \frac{1}{N m_Q} \mathbb{E}_Q \left[\int_0^{\tau_N} 1_A\big(\theta^s \kappa(Y)\big)\, ds \right], \qquad N \in \mathbb{N}. \tag{B.29}
$$

Suppose that Q is also ergodic. Then $\lim_{N\to\infty} \tau_N / N = m_Q$ Q-a.s. and in $L^1(Q)$. Hence, for given $\varepsilon > 0$ we can find a $T_0(\varepsilon)$ such that, for $T \geq T_0(\varepsilon)$,

$$
\mathbb{E}_Q \Big[\Big| \frac{\tau_{N(T)} - T}{m_Q N(T)} \Big| \Big] + \Big| \frac{T}{m_Q N(T)} - 1 \Big| \leq \varepsilon, \tag{B.30}
$$

where $N(T)=\lceil T/m_Q\rceil$. Thus, for $T\geq T_0(\varepsilon)$ and any measurable $A\subset C[0,\infty)$, we have

$$\begin{aligned}&\Big|\Psi_Q(A)-\frac{1}{T}\int_0^T\big(\kappa(Q)\circ(\theta^s)^{-1}\big)(A)\,ds\Big|\\&\leq\frac{1}{m_QN(T)}\Big|\mathbb{E}_Q\Big[\int_0^{\tau_{N(T)}}1_A\big(\theta^s\kappa(Y)\big)\,ds-\int_0^T1_A\big(\theta^s\kappa(Y)\big)\,ds\Big]\Big|\\&\quad+\Big|\Big(\frac{1}{m_QN(T)}-\frac{1}{T}\Big)\int_0^T1_A\big(\theta^s\kappa(Y)\big)\,ds\Big|\leq\mathbb{E}_Q\Big[\Big|\frac{\tau_{N(T)}-T}{m_QN(T)}\Big|\Big]+\Big|\frac{T}{m_QN(T)}-1\Big|\leq\varepsilon,\end{aligned}\tag{B.31}$$

i.e., (B.28) holds.

If Q is not ergodic, then use the ergodic decomposition

$$Q=\int_{\mathcal{P}^{\mathrm{erg,fin}}(F^{\mathbb{N}})}Q'\,W_Q(dQ')\tag{B.32}$$

and note that

$$m_Q=\int_{\mathcal{P}^{\mathrm{erg,fin}}(F^{\mathbb{N}})}m_{Q'}\,W_Q(dQ'),\quad\Psi_Q=\int_{\mathcal{P}^{\mathrm{erg,fin}}(F^{\mathbb{N}})}\frac{m_{Q'}}{m_Q}\,\Psi_{Q'}\,W_Q(dQ')\tag{B.33}$$

(see also [2, Section 6]). We can choose $N(T)$ so large that the set of Q's for which (B.30) holds (with Q replaced by Q') has W_Q-measure arbitrarily close to 1. □

B.3 Continuity of the "letter part" of the rate function under truncation: discrete-time

In this section we consider a discrete-time scenario as in [2]: $\rho\in\mathcal{P}(\mathbb{N})$, E is a Polish space, $\nu\in\mathcal{P}(E)$, the sequence of words $(Y^{(i)})_{i\in\mathbb{N}}$ with discrete lengths has reference law $q_{\rho,\nu}^{\otimes\mathbb{N}}$ with $q_{\rho,\nu}$ as in [2, Eq. (1.4)]. The following lemma extends [2, Lemma A.1] to Polish spaces (in [2] it was only proved and used for finite E, and without explicit control of the error term). Via coarse-graining, this lemma was used in the proof of Proposition 2.5.

Lemma B.5. *Let $Q\in\mathcal{P}^{\mathrm{fin}}(\widetilde{E}^{\mathbb{N}})$ and $0<\varepsilon<\frac12$. Let $\mathrm{tr}\in\mathbb{N}$ be so large that*

$$\mathbb{E}_Q\Big[\big(|Y^{(1)}|-\mathrm{tr}\big)_+\Big]<\frac{\varepsilon}{2}m_Q.\tag{B.34}$$

Then

$$(1-\varepsilon)\big(H(\Psi_{[Q]_{\mathrm{tr}}}\mid\nu^{\otimes\mathbb{N}})+b(\varepsilon)\big)\leq H(\Psi_Q\mid\nu^{\otimes\mathbb{N}})\tag{B.35}$$

with $b(\varepsilon)=-2\varepsilon+[\varepsilon\log\varepsilon+(1-\varepsilon)\log(1-\varepsilon)]/(1-\varepsilon)$, satisfying $\lim_{\varepsilon\downarrow0}b(\varepsilon)=0$. In particular,

$$\lim_{\mathrm{tr}\to\infty}H(\Psi_{[Q]_{\mathrm{tr}}}\mid\nu^{\otimes\mathbb{N}})=H(\Psi_Q\mid\nu^{\otimes\mathbb{N}}).\tag{B.36}$$

Proof. We can assume w.l.o.g. that $H(\Psi_Q\mid\nu^{\otimes\mathbb{N}})<\infty$ for otherwise (B.35) is trivial and (B.36) follows from lower-semicontinuity of specific relative entropy.

First, assume that Q is ergodic, then Ψ_Q is ergodic as well (see [1, Remark 5]). For $\Psi\in\mathcal{P}^{\mathrm{erg}}(E^{\mathbb{N}})$ and $\delta\in(0,1)$,

$$\begin{aligned}H(\Psi\mid\nu^{\otimes\mathbb{N}})&=\lim_{L\to\infty}-\frac{1}{L}\log\Big(\inf\big\{\nu^{\otimes L}(B)\colon\ B\subset E^L,(\pi_L\Psi)(B)\geq1-\delta\big\}\Big),&\text{(B.37)}\\&=\lim_{L\to\infty}\sup\Big\{-\frac{1}{L}\log\nu^{\otimes L}(B):\ B\subset E^L,(\pi_L\Psi)(B)\geq1-\delta\Big\}.&\text{(B.38)}\end{aligned}$$

This replaces the asymptotics of the covering number and its relation to specific entropy for ergodic measures on discrete shift spaces that was employed in the proof of [2, Lemma A.1], and can be deduced with bare hands from the Shannon-McMillan-Breiman theorem. Indeed, asymptotically optimal B's are of the form $\{\frac{1}{L}\log\frac{d\pi_L\Psi}{d\nu^{\otimes L}} \in H(\Psi \mid \nu^{\otimes\mathbb{N}}) \pm \epsilon\}$: Put $f_L = \frac{d\pi_L\Psi}{d\nu^{\otimes L}}$ and set $B_L = \{\frac{1}{L}\log f_L > H(\Psi \mid \nu^{\otimes\mathbb{N}}) - \epsilon\}$. Then $(\pi_L\Psi)(B_L) \to 1$ by the Shannon-McMillan-Breiman, and $\nu^{\otimes L}(B_L) = \int_{B_L}\frac{1}{f_L}d\pi_L\Psi \le \exp[-L(H(\Psi \mid \nu^{\otimes\mathbb{N}}) - \epsilon)]$, i.e., the right-hand side of (B.38) is $\ge H(\Psi \mid \nu^{\otimes\mathbb{N}})$. For the reverse inequality, consider any $B \subset E^L$ with $(\pi_L\Psi)(B) \ge \frac{1}{2}$, say. Set $B' = B \cap \{\frac{1}{L}\log f_L < H(\Psi \mid \nu^{\otimes\mathbb{N}}) + \epsilon\}$. Then $\pi_L\Psi(B') \ge \frac{1}{3}$ for L large enough and $\nu^{\otimes L}(B) \ge \nu^{\otimes L}(B') \ge \exp[-L(H(\Psi \mid \nu^{\otimes\mathbb{N}}) + \epsilon)]\pi_L\Psi(B')$. Hence the right-hand side of (B.38) is also $\le H(\Psi \mid \nu^{\otimes\mathbb{N}})$.

To check (B.35), fix $\varepsilon > 0$. For L sufficiently large, we construct a set $B_L \subset E^L$ such that $\pi_L\Psi_Q(B_L) \ge \frac{1}{2}$ and $\nu^{\otimes L}(B_L) \le \exp[-L(1-\varepsilon)(b_L(\varepsilon) + H(\Psi_{[Q]_{\mathrm{tr}}} \mid \nu^{\otimes\mathbb{N}}))]$, i.e.,

$$-\frac{1}{L}\log\nu^{\otimes L}(B_L) \ge (1-\varepsilon)\big[H(\Psi_{[Q]_{\mathrm{tr}}} \mid \nu^{\otimes\mathbb{N}}) + b_L(\varepsilon)\big], \tag{B.39}$$

where $\lim_{L\to\infty} b_L(\varepsilon) = b(\varepsilon)$. Via (B.38) applied to $\Psi = \Psi_Q$, this yields (B.35).

To construct the sets B_L, we proceed as follows. Put $N = \lceil(1+2\varepsilon)L/m_Q\rceil$. By the ergodicity of Q (see [2, Section 3.1] for analogous arguments), we can find a set $A \subset \widetilde{E}^N$ such that

$$\forall\,(y^{(1)},\dots,y^{(N)}) \in A\colon$$

$$|\kappa(y^{(1)},\dots,y^{(N)})| \ge L(1+\varepsilon), \quad |y^{(1)}| \le \mathrm{tr}, \quad \sum_{i=1}^{N}(|y^{(i)}| - \mathrm{tr})_+ < \varepsilon L, \tag{B.40}$$

$$\mathbb{E}_Q\Big[|Y^{(1)}|1_A(Y^{(1)},\dots,Y^{(N)})\Big] \ge (1-\varepsilon)m_Q, \tag{B.41}$$

and the set

$$B'_L = B'_L(A) = \Big\{\pi_L\big(\theta^i\kappa([y^{(1)}]_{\mathrm{tr}},\dots,[y^{(N)}]_{\mathrm{tr}})\big)\colon (y^{(1)},\dots,y^{(N)}) \in A, i = 0,1,\dots,|y^{(1)}|-1\Big\} \subset E^L \tag{B.42}$$

satisfies

$$\pi_L\Psi_{[Q]_{\mathrm{tr}}}(B'_L) \ge \frac{1}{2}, \quad \nu^{\otimes\lceil L(1-\varepsilon)\rceil}(\pi_{\lceil L(1-\varepsilon)\rceil}B'_L) \le \exp\big[-L(1-\varepsilon)\big(H(\Psi_{[Q]_{\mathrm{tr}}} \mid \nu^{\otimes\mathbb{N}}) - 2\varepsilon\big)\big]. \tag{B.43}$$

Here, use (B.34) in (B.40–B.41), and note that $N\big(1-\frac{\varepsilon}{2}\big)m_Q \sim (1+2\varepsilon)\big(1-\frac{\varepsilon}{2}\big)L \ge (1+\varepsilon)L$ and $N\frac{\varepsilon}{2}m_Q \sim (1+2\varepsilon)\frac{\varepsilon}{2}L < \varepsilon L$ as $L \to \infty$.

For $I \subset \{1,\dots,L\}$, $x \in E^L$ and $y \in E^{|I|}$, write $\mathsf{ins}_I(x;y) \in E^{L+|I|}$ for the word of length $L+|I|$ consisting of the letters from y at index positions in I and the letters from x at index positions not in I, with the order of letters preserved within x and within y (the word y is inserted in x at the positions in I). Put

$$B_L = \pi_L\Big(\big\{\mathsf{ins}_I(x;y)\colon\ x \in B'_L, I \subset \{1,\dots,L\}, |I| \le \varepsilon L, y \in E^{|I|}\big\}\Big). \tag{B.44}$$

Then $\pi_L\Psi_Q(B_L) \ge \frac{1}{2}$ by construction. Furthermore, for fixed $I \subset \{1,\dots,L\}$ with $|I| = k \le \varepsilon L$,

$$\nu^{\otimes L}\Big(\pi_L\big(\big\{\mathsf{ins}_I(x;y)\colon\ x \in B'_L, y \in E^k\big\}\big)\Big) = \nu^{\otimes L}\big(\pi_{L-k}(B'_L)\big) \le \nu^{\otimes\lceil L(1-\varepsilon\rceil]}(\pi_{\lceil L(1-\varepsilon)\rceil}B'_L), \tag{B.45}$$

and hence

$$\nu^{\otimes L}(B_L) \le [\varepsilon L]\binom{L}{[\varepsilon L]} \exp\big[-L(1-\varepsilon)\big(H(\Psi_{[Q]_{\rm tr}} \mid \nu^{\otimes\mathbb{N}}) - 2\varepsilon\big)\big]$$
$$= \exp\big[-L(1-\varepsilon)\big(b_L(\varepsilon) + H(\Psi_{[Q]_{\rm tr}} \mid \nu^{\otimes\mathbb{N}})\big)\big] \tag{B.46}$$

with $b_L(\varepsilon) = -\frac{1}{(1-\varepsilon)L}(\log[\varepsilon L] + \log\binom{L}{[\varepsilon L]}) - 2\varepsilon$, which satisfies $\lim_{\varepsilon\downarrow 0} b_L(\varepsilon) = b(\varepsilon)$.

It remains to prove (B.36). Since $\text{w-lim}_{\rm tr\to\infty} \Psi_{[Q]_{\rm tr}} = \Psi_Q$, we have $\liminf_{\rm tr\to\infty} H(\Psi_{[Q]_{\rm tr}} \mid \nu^{\otimes\mathbb{N}}) \ge H(\Psi_Q \mid \nu^{\otimes\mathbb{N}})$, while the reverse inequality $\limsup_{\rm tr\to\infty} H(\Psi_{[Q]_{\rm tr}} \mid \nu^{\otimes\mathbb{N}}) \le H(\Psi_Q \mid \nu^{\otimes\mathbb{N}})$ follows from (B.34–B.35) and the fact that $\lim_{\rm tr\to\infty} \mathbb{E}_Q[(|Y^{(1)}| - {\rm tr})_+] = m_Q$ by dominated convergence.

For non-ergodic Q, decompose as in [2, Eqs.(6.1)–(6.3)], use the above argument on each of the ergodic components, and use the fact that specific relative entropy is affine. □

C Existence of specific relative entropy

In this section we prove (B.5). For technical reasons, we consider the two-sided scenario. The argument is standard, but the fact that time is continuous requires us to take care.

Proof. Let $\Omega = \tilde{C}(\mathbb{R})$ be the set of continuous functions $\omega\colon \mathbb{R}\to\mathbb{R}$ with $\omega(0) = 0$, which is a Polish space e.g. via the metric $d(\omega,\omega') = \int_{\mathbb{R}} e^{-|t|}\big(|\omega(t)-\omega'(t)| \wedge 1\big)dt$. The shifts on Ω are $\theta^t\omega(\cdot) = \omega(\cdot + t) - \omega(t)$. A probability measure Ψ on Ω has stationary increments when $\Psi = \Psi\circ(\theta^t)^{-1}$ for all $t\in\mathbb{R}$. For an interval $I\subset\mathbb{R}$ denote $\mathcal{F}_I = \sigma(\omega(t)-\omega(s)\colon s,t\in I)$. Ψ_I denotes Ψ restricted to $\mathcal{F}_I$. Write $\mathscr{W}$ for the Wiener measure on Ω, i.e., the law of a (two-sided) Brownian motion.

Let $\Psi\in\mathcal{P}(\Omega)$ with stationary increments be given and assume that $h(\Psi_{[0,T]} \mid \mathscr{W}_{[0,T]}) < \infty$ for all $T>0$. To verify (B.5), we imitate well-known arguments from the discrete-time setup (see e.g. Ellis [8, Section IX.2]).

For I_1, I_2 disjoint intervals in $\mathbb{R}$, denote by $\kappa^\Psi_{I_1,I_2}\colon \Omega\times\mathcal{F}_{I_2}\to[0,1]$ a regular version of the conditional law of (the increments of) Ψ on I_2, given the increments in I_1, i.e., for fixed ω, $\kappa^\Psi_{I_1,I_2}(\omega,\cdot)$ is a probability measure on $\mathcal{F}_{I_2}$, for fixed $A\in\mathcal{F}_{I_2}$, $\kappa^\Psi_{I_1,I_2}(\cdot,A)$ is an $\mathcal{F}_{I_1}$-measurable function, and $\kappa^\Psi_{I_1,I_2}(\omega,A)$ is a version of $\mathbb{E}_\Psi[1_A|\mathcal{F}_{I_1}]$. When $I_1=\emptyset$, $\kappa^\Psi_{\emptyset,I_2}(\omega,A) = \Psi_{I_2}(A)$. Similarly, define $\kappa^{\mathscr{W}}_{I_1,I_2}$ (which is simply $\kappa^{\mathscr{W}}_{I_1,I_2}(\omega,A) = \mathscr{W}_{I_2}(A)$ by the independence of the Brownian increments).

Put

$$a_{I_1,I_2} = \int_\Omega \Psi(d\omega_1)\int_\Omega \kappa^\Psi_{I_1,I_2}(\omega_1,d\omega_2)\,\log\left[\frac{d\kappa^\Psi_{I_1,I_2}(\omega_1,\cdot)}{d\kappa^{\mathscr{W}}_{I_1,I_2}(\omega_1,\cdot)}(\omega_2)\right], \tag{C.1}$$

the expected relative entropy of the conditional distribution under Ψ on $\mathcal{F}_{I_2}$ given $\mathcal{F}_{I_1}$ w.r.t. Wiener measure on $\mathcal{F}_{I_2}$). We have $a_{I_1,I_2}<\infty$ for bounded intervals, because of the assumption of finite relative entropy of Ψ w.r.t. $\mathscr{W}$ on compact time intervals. By stationarity, $a_{I_1,I_2} = a_{t+I_1,t+I_2}$ for any t.

Let $I_1'\subset I_1$, note that $\kappa^\Psi_{I_1,I_2}(\omega,\cdot) \ll \kappa^\Psi_{I_1',I_2}(\omega,\cdot)$ for Ψ-a.e. ω, and $\kappa^{\mathscr{W}}_{I_1,I_2}(\omega,\cdot) = \kappa^{\mathscr{W}}_{I_1',I_2}(\omega,\cdot) = \mathscr{W}_{I_2}(\cdot)$. By the consistency property of conditional distributions, we have

$$a_{I_1',I_2} = \int_\Omega \Psi(d\omega_1)\int_\Omega \kappa^\Psi_{I_1,I_2}(\omega_1,d\omega_2)\log\left[\frac{d\kappa^\Psi_{I_1',I_2}(\omega_1,\cdot)}{d\kappa^{\mathscr{W}}_{I_1',I_2}(\omega_1,\cdot)}(\omega_2)\right]. \tag{C.2}$$

Indeed,

$$\int_\Omega \Psi(d\omega_1) \int_\Omega \kappa^\Psi_{I_1,I_2}(\omega_1, d\omega_2) f(\omega_1,\omega_2) = \int_\Omega \Psi(d\omega_1) \int_\Omega \kappa^\Psi_{I_1',I_2}(\omega_1, d\omega_2) f(\omega_1,\omega_2) \tag{C.3}$$

for any function $f(\omega_1,\omega_2)$ that is $\mathcal{F}_{I_1'} \otimes \mathcal{F}_{\mathbb{R}}$-measurable. Hence

$$\begin{aligned} & a_{I_1,I_2} - a_{I_1',I_2} \qquad\qquad (C.4)\\ &= \int_\Omega \Psi(d\omega_1) \int_\Omega \kappa^\Psi_{I_1,I_2}(\omega_1, d\omega_2) \Bigg(\log\left[\frac{d\kappa^\Psi_{I_1,I_2}(\omega_1,\cdot)}{d\kappa^{\mathscr{W}}_{I_1,I_2}(\omega_1,\cdot)}(\omega_2)\right] - \log\left[\frac{d\kappa^\Psi_{I_1',I_2}(\omega_1,\cdot)}{d\kappa^{\mathscr{W}}_{I_1',I_2}(\omega_1,\cdot)}(\omega_2)\right]\Bigg) \\ &= \int_\Omega \Psi(d\omega_1) \int_\Omega \kappa^\Psi_{I_1,I_2}(\omega_1, d\omega_2) \log\left[\frac{d\kappa^\Psi_{I_1,I_2}(\omega_1,\cdot)}{d\kappa^{\Psi}_{I_1',I_2}(\omega_1,\cdot)}(\omega_2)\right] \geq 0 \qquad\qquad (C.5)\end{aligned}$$

because the inner integral is $h(\kappa^\Psi_{I_1,I_2}(\omega_1,\cdot) \mid \kappa^\Psi_{I_1',I_2}(\omega_1,\cdot)) \geq 0$. Choosing $I_1' = \emptyset$, (C.5), we get $a_{I_1,I_2} \geq a_{\emptyset,I_2} = h(\Psi_{I_2} \mid \mathscr{W}_{I_2})$.

Observe

$$\frac{d\Psi_{(0,s+t]}}{d\mathscr{W}_{(0,s+t]}}(\omega) = \frac{d\Psi_{(0,t]}}{d\mathscr{W}_{(0,t]}}(\omega)\, \frac{d\kappa^\Psi_{(0,t],(t,s+t]}(\omega,\cdot)}{d\kappa^{\mathscr{W}}_{(0,t],(t,s+t]}(\omega,\cdot)}(\omega) \quad \Psi_{(0,s+t]} - \text{a.s.}, \tag{C.6}$$

take logarithms and integrate w.r.t. Ψ (using consistency of conditional expectation on the right-hand side), to obtain

$$h\big(\Psi_{(0,s+t]} \mid \mathscr{W}_{(0,s+t]}\big) = h\big(\Psi_{(0,t]} \mid \mathscr{W}_{(0,t]}\big) + a_{(0,t],(t,s+t]} \geq h\big(\Psi_{(0,t]} \mid \mathscr{W}_{(0,t]}\big) + h\big(\Psi_{(0,s]} \mid \mathscr{W}_{(0,s]}\big). \tag{C.7}$$

Thus, the function $(0,\infty) \ni t \mapsto h(\Psi_{(0,t]} \mid \mathscr{W}_{(0,t]})$ is super-additive, and (B.5) follows from Fekete's lemma. □

Under $\kappa^\Psi_{(-\infty,0],(0,h]}$, the coordinate process will be a Brownian motion with a (possibly complicated) drift process $U_t = \int_0^t u_s\, ds$, where $(u_t)_{t\geq 0}$ can be chosen adapted, and

$$\mathbb{E}_\Psi\big[h(\kappa^\Psi_{(-\infty,0],(0,h]} \mid \mathscr{W}_{(0,h]})\big] = \mathbb{E}_\Psi\big[\textstyle\int_0^h u_s^2\, ds\big] \tag{C.8}$$

(see Föllmer [10]).

References

[1] M. Birkner, Conditional large deviations for a sequence of words, Stoch. Proc. Appl. 118 (2008), 703–729.

[2] M. Birkner, A. Greven and F. den Hollander, Quenched large deviation principle for words in a letter sequence, Probab. Theory Relat. Fields 148 (2010), 403–456.

[3] M. Birkner and F. den Hollander, A variational view on the Brownian copolymer, manuscript in preparation.

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

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

[6] F. Caravenna, G. Giacomin and F.L. Toninelli, Copolymers at selective interfaces: settled issues and open problems, In: *Probability in Complex Physical Systems. In honour of Erwin Bolthausen and Jürgen Gärtner* (eds. J.-D. Deuschel, B. Gentz, W. König, M. von Renesse, M. Scheutzow, U. Schmock), Springer Proceedings in Mathematics 11, 2012, pp. 289–312.

[7] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications* (2nd. Ed.), Springer, 1998.

[8] R. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, Springer, 1985.

[9] S. Ethier and T. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, 1986.

[10] H. Föllmer, Time reversal on Wiener space, Lecture Notes in Math. 1158, Springer, 1986, pp. 119–129.

[11] H. Föllmer, Random fields and diffusion processes, in: École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87, pp. 101–203, Lecture Notes in Math. 1362, Springer, 1988.

[12] R.M. Gray, *Entropy and Information Theory*, Springer, 1991.
`http://ee.stanford.edu/~gray/it.html`

[13] C. Kipnis and C. Landim, *Scaling Limits of Interacting Particle Systems*, Springer, 1999.