Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Quenched large deviations for simple random walks on percolation clusters including long-range correlations

Noam Berger[1,2], Chiranjib Mukherjee[3], Kazuki Okamura[4]

submitted: December 22, 2016

[1] The Hebrew University of Jerusalem
Einstein Institute of Mathematics
Edmond J. Safra Campus, Givat Ram
Jerusalem, 91904
Israel
E-Mail: noam.berger@tum.de

[2] Technische Universität München
Lehrstuhl für Wahrscheinlichkeitstheorie
Boltzmannstr. 3
85748 Garching
Germany

[3] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: chiranjib.mukherjee@wias-berlin.de

[4] Kyoto University
Research Institute for Mathematical Sciences
Kyoto, 606-8502
Japan
E-Mail: kazukio@kurims.kyoto-u.ac.jp

No. 2360

Berlin 2016

2010 *Mathematics Subject Classification.* 60J65, 60J55, 60F10, 60K37.

Key words and phrases. Large deviations, random walk on percolation clusters, long-range correlations, random interlacements, Gaussian free field, random cluster model.

Most of the work in this paper was carried out when the second author was a visiting assistant professor at the Courant Institute of Mathematical Sciences, New York University in the academic year 2015-2016, and its hospitality is gratefully acknowledged. The authors would like to thank S.R.S.Varadhan for reading an early draft of the manuscript and many valuable suggestions. The third author was supported by Grant-in-Aid for Research Activity Start-up (15H06311) and Grant-in-Aid for JSPS Fellows (16J04213).

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

ABSTRACT. We prove a quenched large deviation principle (LDP)for a simple random walk on a supercritical percolation cluster (SRWPC) on the lattice.The models under interest include classical Bernoulli bond and site percolation as well as models that exhibit long range correlations, like the random cluster model, the random interlacement and its vacant set and the level sets of the Gaussian free field.

Inspired by the methods developed by Kosygina, Rezakhanlou and Varadhan ([KRV06]) for proving quenched LDP for elliptic diffusions with a random drift, and by Yilmaz ([Y08]) and Rosenbluth ([R06]) for similar results regarding elliptic random walks in random environment, we take the point of view of the moving particle and prove a large deviation principle for the quenched distribution of the pair empirical measures if the environment Markov chain in the non-elliptic case of SRWPC. Via a contraction principle, this reduces easily to a quenched LDP for the distribution of the mean velocity of the random walk and both rate functions admit explicit variational formulas.

The main approach of our proofs are based on exploiting coercivity properties of the relative entropy in the context of convex variational analysis, combined with input from ergodic theory and invoking geometric properties of the percolation cluster under supercriticality.

1. MOTIVATION, INTRODUCTION AND MAIN RESULTS

We consider a simple random walk on the infinite cluster of some bond and site percolation models on $\mathbb{Z}^d$, $d \geq 2$. The percolation models under interest include classical Bernoulli bond and site percolation, as well as models that exhibit long-range correlations, including the random-cluster model, random interlacements and its vacant set in $d \geq 3$, and the level set of the Gaussian free field (also for $d \geq 3$). Conditional on the event that the origin lies in the infinite open cluster, it is known that a law of large numbers and quenched central limit theorem hold (see [SS04], [MP07], [BB07] and [PRS15]). Treatment of these classical questions for these models need care because of its inherent *non-ellipticity* – a problem which permeates in several forms in the above mentioned literature.

Questions on large deviation principles (LDP) in the quenched setting for general random walks in *elliptic* random environments (RWRE) have also been studied. In $d = 1$, first Greven and den Hollander ([GdH98]) for i.i d. and uniformly elliptic random environments, and then Comets, Gantert and Zeitouni ([CGZ00]) for stationary, ergodic and uniformly elliptic random environments, derived quenched LDP for the mean velocity of a a RWRE and obtained explicit variational formulas for the rate function. For $d \geq 1$, Zerner ([Z98], see also Sznitman ([S94])) proved quenched LDP under the assumption that the logarithm of the random walk transition probabilities possesses finite d-th moment and the random environment enjoys the *nestling property*. His method is based on proving shape theorems invoking the sub-additive ergodic theorem. Using the sub-additivity more directly, Varadhan ([V03]) proved a quenched LDP dropping the *nestling* assumption and assuming uniform ellipticity for the random environment. However, the use of sub-additivity in the above results did not lead to any desired formula for the rate function.

Kosygina, Rezakhanlou and Varadhan ([KRV06]) derived a novel method for proving quenched LDP using the *environment seen from the particle* in the context of a diffusion with a random drift assuming some growth conditions on the random drift (ellipticity) and obtained a variational formula for the rate function. This method goes parallel to quenched homogenization of random Hamilton-Jacobi-Bellman (HJB) equations. Rosenbluth ([R06]) adapted this theory to the “level-1"large deviation analysis of the rescaled location of a multidimensional random walk in random environments and also obtained a formula for the rate function. The assumption regarding the growth condition on the random drift imposed in [KRV06] under which homogenization of HJB takes place, or quenched large deviation principle for the rescaled law of the diffusion holds, now translates to the assumption that logarithm of the random walk transition probabilities possesses finite $d + \varepsilon$ moment, for some $\varepsilon > 0$ (see [R06]). Under the same moment assumption, Yilmaz ([Y08]) extended this work to a “level-2"LDP for the law of the pair empirical measures of the environment Markov chain and subsequently Rassoul-Agha and Seppäläinen ([RS11]) proved a “level-3"LDP for the empirical process for the environment Markov chain. Like Rosenbluth ([R06]), both [Y08] and [RS11] obtained variational formulas for the corresponding rate functions. This method has been further exploited for studying free energy for directed and non-directed random walks in a unbounded random potential (see the works of Rassoul-Agha, Seppäläinen and Yilmaz [RSY13, RSY14] and Georgiou et al. [GRSY13, GRS14]). We also refer to the works of Armstrong and Souganidis ([AS12]) for the continuous analogue of [RSY13] concerning homogenization of random Hamilton Jacobi Bellman equations in unbounded environments. Roughly speaking, all these results in the aforementioned literature work only under the assumption that $V := -\log \pi \in L^p(\mathbb{P})$ with $p > d$, where π denotes the random walk transition probabilities in the elliptic random environment whose law is denoted by $\mathbb{P}$. Thus, the aforementioned literature does not cover the case $V = \infty$ pertinent to the case of a random walk on a supercritical percolation cluster, an important model that carries the aforementioned inherent non-ellipticity of the random environment.

In this context, it is the goal of the present article to study quenched large deviation principles for the distribution of the empirical measures of the environment Markov chain of SRWPC (*level-2*) and subsequently deduce the particle dynamics of the rescaled location (*level-1*) of the walk on the cluster. We start with a precise mathematical layout of the random environments under interest including the bond and site percolations on $\mathbb{Z}^d$.

1.1 The bond percolation models.

We fix $d \geq 2$ and denote by $\mathbb{B}_d$ the set of nearest neighbor edges of the lattice $\mathbb{Z}^d$ and by $\mathbb{U}_d = \{\mathfrak{p}e_i\}_{i=1}^d$ the set of edges from the origin to its nearest neighbor. Let $\Omega = \{0,1\}^{\mathbb{B}_d}$ be the space of all *percolation configurations* $\omega = (\omega_b)_{b \in \mathbb{B}_d}$. In other words, $\omega_b = 1$ refers to the edge $0 \leftrightarrow b$ being *present* or *open*, while $\omega_b = 0$ implies that it is *vacant* or *closed*. Let $\mathcal{B}$ be the Borel-σ-algebra on Ω defined by the product topology. Note that $\mathbb{Z}^d$ acts as a group on $(\Omega, \mathcal{B})$ via translations. In other words, for each $x \in \mathbb{Z}^d$, $\tau_x : \Omega \longrightarrow \Omega$ acts as a *shift* given by $(\tau_x \omega)_b = \omega_{x+b}$. Let $\mathbb{P}$ be a probability measure on Ω. Let $\Omega_0 = \{\omega \colon 0 \in \mathcal{C}_\infty(\omega)\}$. We define the conditional probability $\mathbb{P}_0$ by

$$\mathbb{P}_0(A) = \mathbb{P}\big(A \big| \Omega_0\big) \qquad A \in \mathcal{B}.$$

• **I.I.D. bond percolation.** We fix the *percolation parameter* $p \in (0,1)$ and denote by

$$\mathbb{P} = \mathbb{P}_p := \big(p\delta_1 + (1-p)\delta_0\big)^{\mathbb{B}_d}$$

the product measure with marginals $\mathbb{P}(\omega_b = 1) = p = 1 - \mathbb{P}(\omega_b = 0)$. Note that the product measure $\mathbb{P}$ is invariant under this action. It is known that there is a critical percolation probability $p_c = p_c(d)$ which is the infimum of all p's such that $\mathbb{P}(0 \in \mathcal{C}_\infty) > 0$. In this paper we only consider the case $p > p_c$. By Burton-Keane's uniqueness theorem ([BK89]), the infinite cluster is unique and so $\mathcal{C}_\infty$ is connected with $\mathbb{P}$-probability one.

• **Random cluster model.** The second example is the random-cluster model, which is a natural extension of Bernoulli bond percolation. However, this models exhibits long range correlations and one necessarily drops the i.i.d. structure present in the first example. Let us shortly recall the basic structure and the salient properties of this model.

Let $d \geq 2$, $p \in [0,1]$, $q \geq 1$, and let also Λ be a box in $\mathbb{Z}^d$ with boundary condition $\xi \in \{0,1\}^{\mathbb{B}_d}$. Let $\mathbb{P}^{\xi}_{\Lambda,p,q}$ be the random-cluster measure on Λ, defined as

$$\mathbb{P}^{\xi}_{\Lambda,p,q}(\{\omega\}) = \frac{1}{Z}\, p^{n(\omega)}\, (1-p)^{|\Lambda| - n(\omega)}\, q^{k(\omega)}.$$

Here Z is a normalizing constant that makes $\mathbb{P}^{\xi}_{\Lambda,p,q}$ a probability measure, while $n(\omega)$ is the number of edges in $\Lambda \cap \omega$, $|\Lambda|$ is the number of all edges in Λ. $k(\omega)$ is the number of open clusters of $\omega_{\Lambda,\xi}$ intersecting Λ, where

$$\omega_{\Lambda,\xi} = \begin{cases} \omega & \text{on } \Lambda \\ \xi & \text{outside } \Lambda. \end{cases}$$

Let

$$\mathbb{P}^{(\mathbf{b})}_{p,q} = \lim_{\Lambda \to \mathbb{Z}^d} \mathbb{P}^{(\mathbf{b})}_{\Lambda,p,q} \qquad \text{where } \mathbf{b} = (b, b, \dots) \in \{0,1\}^{\mathbb{B}_d},\ b \in \{0,1\}.$$

In other words, $\mathbb{P}^{(\mathbf{b})}_{p,q}$ is the extremal infinite-volume limit random-cluster measures, with free (for $\mathbf{b} = 0$) and wired (for $\mathbf{b} = 1$) conditions respectively. For any $b \in \{0,1\}$, let

$$p_c^{(b)}(q) = \inf\Big\{p \in [0,1] : \mathbb{P}^{(b)}_{p,q}(0 \leftrightarrow \infty) > 0\Big\}, \quad b = 0, 1.$$

Then, $p_c^{(0)}(q) = p_c^{(1)}(q) \in (0,1)$ and we write this as $p_c(q)$. It is well-known that, for both $b=0$ and $b=1$, the measure $\mathbb{P} := \mathbb{P}^{(\mathbf{b})}_{p,q}$ is invariant and ergodic with respect to τ_x for any $x \in \mathbb{Z}^d \setminus \{0\}$ and for all $p \in [0,1]$ ([G06, (4.19) and (4.23)]). Furthermore, for any $p > p_c(q)$, there exists a unique infinite cluster $\mathcal{C}_\infty$, $\mathbb{P}^b_{p,q}$-a.s. by [G06, Theorem 5.99],

For our purpose, we also need the notion of *slab critical probability*, which is defined as follows. For $d \geq 3$, we let

$$\begin{aligned} &S(L,n) := [0, L-1] \times [-n,n]^{d-1} \\ &\widehat{p}_c(q,L) := \inf\left\{p : \liminf_{n\to\infty} \inf_{x\in S(L,n)} \mathbb{P}^{(\mathbf{0})}_{S(L,n),p,q}(0 \leftrightarrow x) > 0\right\} \\ &\widehat{p}_c(q) := \lim_{L\to\infty} \widehat{p}_c(q,L). \end{aligned} \tag{1.1}$$

For $d=2$, for $e_n = (n,0) \in \mathbb{R}^2$, we let

$$\begin{aligned} p_c(q) &:= \sup\left\{p : \lim_{n\to\infty} \frac{-\log \mathbb{P}^{(\mathbf{0})}_{p,q}(0 \leftrightarrow e_n)}{n} > 0\right\}, \\ \widehat{p}_c(q) &:= \frac{q(1-p_g(q))}{p_g(q) + q(1-p_g(q))} \end{aligned} \tag{1.2}$$

and we have the bound $\widehat{p}_c(q) \geq p_c(q)$. Although equality is believed to be true in the last implication ([G06, Conjecture 5.103]), to the best of our knowledge, the only known proofs are available only for the case $q=1$ (i.e., the case of Bernoulli bond percolation (see Grimmett and Marstrand [GM90]), and for $d=2$ and any $q \geq 1$ (see Beffara and Duminil-Copin [BD12]), and for $d \geq 3$ and $q=2$ (i.e., *FK-Ising model*, see Bodineau [B05]). We will henceforth work in the regime that

$$p > \widehat{p}_c(q),$$

and will write $\mathbb{P} = \mathbb{P}^{(\mathbf{b})}_{p,q}$ and $\mathbb{P}_0 = \mathbb{P}(\cdot \,|\, 0 \in \mathcal{C}_\infty)$ throughout the rest of the article.

We point out that in the process of proving our main results (stated in Section 2) corresponding to the random cluster model, we prove some geometric properties of this model as a necessary by-product. In particular, we prove a “chemical distance estimate"between two points in the infinite cluster $\mathcal{C}_\infty$ (see Lemma 4.2), and also obtain exponential tail bounds for the graph distance between the origin and the “first arrivalöf the infinite cluster $\mathcal{C}_\infty$ on any coordinate direction (see Lemma 4.3). Although both results are part of the standard folklore in the i.i.d. percolation literature, the proofs of these two assertions for the random cluster model seem to be new, to the best of our knowledge.

1.2 Site percolation models. The second class of models we are interested in concerns *site percolations*, which include the classical Bernoulli i.i.d. percolation as well as models that carry long-range correlation. We turn to short descriptions of these models.

• **Random interlacements and its vacant set in $d \geq 3$.** This model was introduced by Sznitman [S10]. Let $\mathbb{T}_N = \big(\mathbb{Z}/N\mathbb{Z}\big)^d$ be the discrete torus in $d \geq 3$. For any $u > 0$, the *random interlacement* $\mathcal{I}^{(u)}$ is defined to be a subset of $\mathbb{Z}^d$ which arises as the local limit, as $N \to \infty$ of the sites visited by a simple random walk in $\mathbb{T}_N$ until time $\lfloor uN \rfloor$. For any finite subset $K \subset \mathbb{Z}^d$ with capacity $\mathrm{cap}(K)$, the distribution of $\mathcal{I}^{(u)}$ is given by

$$\mathbb{P}\big[\mathcal{I}^{(u)} \cap K = \emptyset\big] = \mathrm{e}^{-u\mathrm{cap}(K)},$$

Furthermore, $\mathbb{P}$-almost surely, the set $\mathcal{I}^{(u)}$ is an infinite connected subset of $\mathbb{Z}^d$ (see (2.21), [S10]), exhibits long rage correlations given by

$$\Big|\mathbb{P}\big[x,y \in \mathcal{I}^{(u)}\big] - \mathbb{P}\big[x \in \mathcal{I}^{(u)}\big]\,\mathbb{P}\big[y \in \mathcal{I}^{(u)}\big]\Big| \sim \big(1 + |x-y|\big)^{2-d}. \tag{1.3}$$

The *vacant set of random interlacements* $\mathcal{V}^{(u)}$ is defined to be the complement of the random interlacement $\mathcal{I}^{(u)}$ at level u, i.e.,

$$\mathcal{V}^{(u)} = \mathbb{Z}^d \setminus \mathcal{I}^{(u)} \qquad \mathbb{P}\big[K \subset \mathcal{V}^{(u)}\big] = \mathrm{e}^{-u\mathrm{cap}(K)}.$$

Furthermore, $\mathcal{V}^{(u)}$ also exhibits polynomially decaying correlation as in (1.3). It is known that in [TW11] there exists $u_\star \in (0,\infty)$ such that almost surely, for any $u > u_\star$, all connected components of $\mathcal{V}^{(u)}$ are finite, while for $u < u_\star$, $\mathcal{V}^{(u)}$ contains an infinite connected component $\mathcal{C}_\infty$, which is unique (see [T09']). Furthermore, geometric properties of the random interlacements and its vacant set have been studied extensively by Drewitz-Rath-Sapozhnikov ([DRS14]) obtaining sharp estimates on the graph distance, assuming that $u \in (0, \overline{u})$ for some $\overline{u} \le u_\star$ ($\overline{u}$ is introduced in [DRS14, Theorem 2.5]). Although it is believed that $\overline{u} = u_\star$, we will henceforth assume that

$$u < \overline{u}.$$

and in this regime, as before, we will write $\mathbb{P}_0 = \mathbb{P}(\cdot|\, 0 \in \mathcal{C}_\infty)$.

• **Gaussian free fields and its level sets in $d \ge 3$.** This model has a strong background in statistical physics (see [LS86],or [S07] for a mathematical survey. The Gaussian free field on $\mathbb{Z}^d$ for $d \ge 3$, is a centered Gaussian field $\varphi = \big(\varphi(x)\big)_{x \in Z^d}$ under the probability measure $\mathbb{P}$ with covariance function

$$\mathbb{E}[\varphi(x)\varphi(y)] = g(x,y) = c_d|x-y|^{2-d},$$

given by the Green function of the simple random walk on $\mathbb{Z}^d$. This leads to long range correlations exhibited by random field φ. For any $h \in \mathbb{R}$, the *excursion set above level* h is defined as

$$E^{\ge h} = \{x \in \mathbb{Z}^d \colon \varphi(x) \ge h\}$$

and it is known that there exists $h_\star \in [0,\infty)$ such that for any $h < h_\star$, $\mathbb{P}$-almost surely, $E^{\ge h}$ contains a unique infinite connected component and for any $h > h_\star$, all the connected components of $E^{\ge h}$ are finite. Like in the case of random interlacements and vacant set of random interlacements, results on the graph distance for the excursion level set $E^{\ge h}$ was also obtained in [DRS14] on the sub-regime $(-\infty, \overline{h})$ for $\overline{h} \le h_\star$. [DRS14, Remark 2.9] conjectures that $\overline{h} = h_\star \in (0,\infty)$ in all $d \ge 3$ and as before, we will also assume that

$$h \in (-\infty, \overline{h}),$$

which guarantees that the level set $E^{\ge h}$ has a unique infinite connected component $\mathcal{C}_\infty$ and as usual, we will write $\Omega_0 = \{0 \in \mathcal{C}_\infty\}$ and will work with the conditional measure

$$\mathbb{P}_0 = \mathbb{P}(\cdot|0 \in \mathcal{C}_\infty).$$

1.3 The simple random walk on the percolation models.

We now define a (discrete time) simple random walk on the unique supercritical percolation cluster $\mathcal{C}_\infty$ corresponding to the percolation models discussed in the last section.

Let a random walk start at the origin and at each unit of time, the walk moves to a nearest neighbor site chosen uniformly at random from the accessible neighbors. More precisely, for each $\omega \in \Omega_0$, $x \in \mathbb{Z}^d$ and $e \in \mathbb{B}_d$, we set

$$\pi_\omega(x,e) = \frac{\mathbb{1}_{\{\omega_e = 1\}} \circ \tau_x}{\sum_{|e'|=1} \mathbb{1}_{\{\omega_{e'}=1\}} \circ \tau_x} \in [0,1], \tag{1.4}$$

and define a simple random walk $X = (X_n)_{n \ge 0}$ as a Markov chain taking values in $\mathbb{Z}^d$ with the transition probabilities

$$\begin{aligned} &P_0^{\pi,\omega}(X_0 = 0) = 1, \\ &P_0^{\pi,\omega}\big(X_{n+1} = x+e \big| X_n = x\big) = \pi_\omega(x,e). \end{aligned} \tag{1.5}$$

This is a canonical way to “put"the Markov chain on the infinite cluster $\mathcal{C}_\infty$. Henceforth, we will refer to this Markov chain as the *simple random walk on the percolation cluster* (SRWPC).

Let us remark that in the expression of $\pi_\omega(x,e)$ as well as $P_0^{\pi,\omega}$ we have apparently used the notation for bond percolation models appearing in Section 1.1. Very similar expression can be used for these objects pertaining to the site percolation models introduced in Section 1.2 too. To alleviate notation, throughout the rest of the article we will continue to write the expressions (1.4) and (1.5) for the transition kernels $\pi_\omega(x,e)$ and transition probabilities $P_0^{\pi,\omega}$ for the SRWPC corresponding to all the percolation models.

2. Main results

In the sequel, in Section 2.1 we will first introduce the environment Markov chain, its empirical measures and certain relative entropy functionals which will be used later. In Section 2.2, we will announce our main results. In Section 2.3 we will carry out a sketch of the existing proof technique related to elliptic RWRE ([Y08],[R06]), comment on the approach taken in the present paper regarding SRWPC and underline the differences to the earlier approach.

2.1 The environment Markov chain. For each $\omega \in \Omega_0$, we consider the process $(\tau_{X_n}\omega)_{n\geq 0}$ which is a Markov chain taking values in the space of environments Ω_0. This is the *environment seen from the particle* and plays an important rôle in the present context, see section 3.1 for a detailed description. We denote by

$$\mathfrak{L}_n = \frac{1}{n}\sum_{n=0}^{n-1} \delta_{\tau_{X_k}\omega, X_{k+1}-X_k} \tag{2.1}$$

the empirical measure of the environment Markov chain and the nearest neighbor steps of the SRWPC $(X_n)_{n\geq 0}$. This is a random element of $\mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)$, the space of probability measures on $\Omega_0 \times \mathbb{B}_d$, which is compact when equipped with the weak topology (note that, $\Omega_0 \subset \Omega$ is closed and hence compact). The empirical measures $\mathfrak{L}_n$ were introduced and their large deviation behavior (in the *quenched setting*) for elliptic random walks in random environments were studied by Yilmaz ([Y08]).

We note that, via the mapping $(\omega,e)\mapsto(\omega,\tau_e\omega)$ the space $\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$ is embedded into $\mathcal{M}_1(\Omega_0\times\Omega)$, and hence, any element $\mu\in\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$ can be thought of as the *pair empirical measure* of the environment Markov chain. In this terminology, we can define its marginal distributions by

$$\begin{aligned} \mathrm{d}(\mu)_1(\omega) &= \sum_{e\in\mathbb{B}_d} \mathrm{d}\mu(\omega,e), \\ \mathrm{d}(\mu)_2(\omega) &= \sum_{e:\ \tau_e\omega'=\omega} \mathrm{d}\mu(\omega',e) = \sum_{e\in\mathbb{B}_d} \mathrm{d}\mu(\tau_{-e}\omega,e). \end{aligned} \tag{2.2}$$

Here $(\mu)_1$ is a measure on Ω_0 and $(\mu)_2$ is a measure on Ω. A relevant subspace of $\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$ is given by

$$\begin{aligned} \mathcal{M}_1^\star = \mathcal{M}_1^\star(\Omega_0\times\mathbb{B}_d) = \Big\{ &\mu\in\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)\colon\ (\mu)_1=(\mu)_2 \ll \mathbb{P}_0 \text{ and } \mathbb{P}_0\text{- almost surely}, \\ &\frac{\mathrm{d}\mu(\omega,e)}{\mathrm{d}(\mu)_1(\omega)} > 0 \text{ if and only if } \omega(e)=1 \text{ for } e\in\mathbb{B}_d \Big\}. \end{aligned} \tag{2.3}$$

We remark that, here $(\mu)_1=(\mu)_2$ means that $(\mu)_2$ is supported on Ω_0 and $(\mu)_1=(\mu)_2$. Furthermore, Lemma 3.1 shows that elements in $\mathcal{M}_1^\star$ are in one-to-one correspondence to Markov kernels (w.r.t. the environment process) on Ω_0 which admit invariant probability measures which are absolutely continuous with respect to $\mathbb{P}_0$.

Finally, we define a *relative entropy functional* $\mathfrak{I}:\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)\to[0,\infty]$ via

$$\mathfrak{I}(\mu) = \begin{cases} \int_{\Omega_0}\sum_{e\in\mathbb{B}_d} \mathrm{d}\mu(\omega,e)\log\frac{\mathrm{d}\mu(\omega,e)}{\mathrm{d}(\mu)_1(\omega)\pi_\omega(0,e)} & \text{if } \mu\in\mathcal{M}_1^\star, \\ \infty & \text{else.} \end{cases} \tag{2.4}$$

For every continuous, bounded and real valued function f on $\Omega_0 \times \mathbb{B}_d$, we denote by

$$\mathfrak{I}^\star(f) = \sup_{\mu \in \mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)} \big\{ \langle f, \mu \rangle - \mathfrak{I}(\mu) \big\}$$

the *Fenchel-Legendre transform* of $\mathfrak{I}(\cdot)$. Likewise, for any $\mu \in \mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)$, $\mathfrak{I}^{\star\star}(\mu)$ denotes the *Fenchel-Legendre transform* of $\mathfrak{I}^\star(\cdot)$.

2.2 Main results: Quenched large deviation principle.

We are now ready to state the main result of this paper, which proves a large deviation principle for the distributions $P_0^{\pi,\omega} \mathfrak{L}_n^{-1}$ on $\mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)$ (usually called *level-2* large deviations) and the distributions $P_0^{\pi,\omega} \frac{X_n}{n}^{-1}$ on $\mathbb{R}^d$ (usually called *level-1* large deviations). Both statements hold true for $\mathbb{P}_0$- almost every $\omega \in \Omega_0$ and in the case of elliptic RWRE, these already exist in the literature (see Yilmaz [Y08] for level-2 large deviations and Rosenbluth [R06] for level-1 large deviations) with the assumption which requires the p-th moment of the logarithm of the RWRE transition probabilities to be finite, for $p > d$. In the present context, due to zero transition probabilities of the SRWPC, we necessarily have to drop this moment assumption.

Before we announce our main result precisely, let us remind the reader that all the percolation models that were introduced in Section 1.1 and Section 1.2 are assumed to be supercritical, the origin is always contained in the unique infinite cluster $\mathcal{C}_\infty$, $\mathbb{P}_0 = \mathbb{P}(\cdot | \{0 \in \mathcal{C}_\infty\})$ denotes the conditional environment measure and $P_0^{\pi,\omega}$ stands for the transition probabilities for SRWPC defined in (1.5). Here is the statement of our first main result.

Theorem 2.1 (Quenched LDP for the pair empirical measures)**.** *Let $d \geq 2$. Then for $\mathbb{P}_0$- almost every $\omega \in \Omega_0$, the distributions of $\mathfrak{L}_n$ under $P_0^{\pi,\omega}$ satisfies a large deviation principle in the space of probability measures on $\mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)$ equipped with the weak topology. The rate function $\mathfrak{I}^{\star\star}$ is the double Fenchel-Legendre transform of the functional $\mathfrak{I}$ defined in* (2.4)*. Furthermore, $\mathfrak{I}^{\star\star}$ is convex and has compact level sets.*

In other words, for every closed set $\mathcal{C} \subset \mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)$, every open set $\mathcal{G} \subset \mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)$ and $\mathbb{P}_0$- almost every $\omega \in \Omega_0$,

$$\limsup_{n\to\infty} \frac{1}{n} \log P_0^{\pi,\omega}\big(\mathfrak{L}_n \in \mathcal{C}\big) \leq -\inf_{\mu \in \mathcal{C}} \mathfrak{I}^{\star\star}(\mu), \tag{2.5}$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \log P_0^{\pi,\omega}\big(\mathfrak{L}_n \in \mathcal{G}\big) \geq -\inf_{\mu \in \mathcal{G}} \mathfrak{I}^{\star\star}(\mu). \tag{2.6}$$

A standard computation shows that the functional $\mathfrak{I}$ defined in (2.4) is convex on $\mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)$. The following lemma, whose proof is based on the “zero speed regimeöf the SRWPC under a supercritical drift and is deferred to until Section 6, shows that $\mathfrak{I}^{\star\star} \neq \mathfrak{I}$.

Lemma 2.2. *Let $d \geq 2$. Then $\mathfrak{I}$ is not lower-semicontinuous on $\mathcal{M}_1(\Omega_0 \times \mathbb{B}_d)$. Hence, $\mathfrak{I} \neq \mathfrak{I}^{\star\star}$.*

We remark that Theorem 2.1 is an easy corollary to the existence of the limit

$$\lim_{n\to\infty} \frac{1}{n} \log E_0^{\pi,\omega}\big\{ \exp\{n\langle f, \mathfrak{L}_n\rangle\}\big\} = \lim_{n\to\infty} \frac{1}{n} \log E_0^{\pi,\omega}\bigg\{ \exp\bigg(\sum_{k=0}^{n-1} f\big(\tau_{X_k}\omega, X_k - X_{k-1}\big)\bigg)\bigg\},$$

for every continuous, bounded function f on $\Omega_0 \times \mathbb{B}_d$ and the symbol $\langle f, \mu\rangle$ denotes, in this context, the integral $\int_{\Omega_0} \mathrm{d}\mathbb{P}_0(\omega) \sum_{e \in \mathbb{B}_d} f(\omega, e) \mathrm{d}\mu(\omega, e)$. We formulate it as a theorem.

Theorem 2.3 (Logarithmic moment generating functions)**.** *For $d \geq 2$, $p > p_c(d)$ and every continuous and bounded function f on $\Omega_0 \times \mathbb{B}_d$,*

$$\lim_{n\to\infty} \frac{1}{n} \log E_0^{\pi,\omega}\bigg\{ \exp\bigg(\sum_{k=0}^{n-1} f\big(\tau_{X_k}\omega, X_k - X_{k-1}\big)\bigg)\bigg\} = \sup_{\mu \in \mathcal{M}_1^\star} \big\{ \langle f, \mu\rangle - \mathfrak{I}(\mu)\big\} \quad \mathbb{P}_0 - \textit{a.s.}$$

We will first prove Theorem 2.3 and deduce Theorem 2.1 directly.

Note that via the contraction map $\xi : \mathcal{M}_1(\Omega_0 \times \mathbb{B}_d) \longrightarrow \mathbb{R}^d$,

$$\mu \mapsto \int_{\Omega_0} \sum_e e \, \mathrm{d}\mu(\omega, e),$$

we have $\xi(\mathfrak{L}_n) = \frac{X_n - X_0}{n} = \frac{X_n}{n}$. Our second main result is the following corollary to Theorem 2.1.

Corollary 2.4 (Quenched LDP for the mean velocity of SRWPC)**.** *Let $d \geq 2$. Then the distributions $P_0^{\pi,\omega}\big(\frac{X_n}{n} \in \cdot\big)$ satisfies a large deviation principle with a rate function*

$$J(x) = \inf_{\mu :\, \xi(\mu)=x} \mathfrak{I}(\mu) \qquad x \in \mathbb{R}^d.$$

Remark 1 Note that Corollary 2.4 has been obtained by Kubota ([K12]) for the SRWPC based on the method of Zerner ([Z98], see also Mourrat ([M12]). Kubota used sub-addtivity and overcame the lack of the moment criterion of Zerner by using classical results about the geometry of the percolation. This way he obtained a rate function which is convex and is the given by the Legendre transform of the *Lyapunov exponents* derived by Zerner ([Z98]). However, using the sub-additive ergodic theorem one does not get any expression or formula for the rate function, nor does the sub-additivity seem amenable for deriving a *level 2* quenched LDP as in Theorem 2.1.

2.3 Survey of earlier proof technique in the elliptic case and comparison with our method.

Earlier relevant work for quenched large deviations was carried out by Kosygina-Rezakhanlou-Varadhan ([KRV06]) for elliptic diffusions in a random drift. Rosenbluth ([R06]) first adapted this approach to the case of elliptic RWRE and derived a level-1 quenched large deviation principle for the distribution of the mean-velocity (the so-called *level-1* large deviations, recall Corollary 2.4). Yilmaz ([Y08]) then extended Rosenbluth's work on elliptic RWRE to a finer large deviation result for the pair empirical measures of the environment Markov chain (the so-called *level-2* large deviations, recall Theorem 2.1). In the present case of deriving similar level-2 quenched large deviations for SRWPC, as a guiding philosophy, we also follow the main steps of Yilmaz ([Y08]). However, due to fundamental obstacles that come up in several facets stemming from the inherent non-ellipticity of the percolation models, an actual execution of the existing method [Y08] fails for the present case of SRWPC. In order to put our present work in context, in this section we will first present a brief survey on the method of Yilmaz ([Y08]), and to emphasize on the similarities and differences of our approach to the earlier one, we will also briefly sketch the main steps for our present technique which allows the treatment of models that are non-elliptic (*and also elliptic*). This will also underline the technical novelty of the present work.

Earlier approach for elliptic RWRE ([Y08]**):** To keep notation consistent, in this survey we will continue to denote by $\mathbb{P}$ the law of a stationary and ergodic random environment and by $\pi(\omega, \cdot)$ we will denote the random walk transition probabilities in the random environment that enjoys the ellipticity (moment) condition $\int |\log \pi|^{d+\varepsilon} \mathrm{d}\mathbb{P} < \infty$ for some $\varepsilon > 0$, as required in [Y08]. The crucial argument is the existence of the limiting logarithmic moment generating function (recall Theorem 2.3) whose proof splits into three main steps:

Lower bound. The lower bound part is based on a classical change of measure argument for the environment Markov chain, followed by an application of an ergodic theorem for the tilted Markov chain. This ergodic theorem is standard (see Kozlov [K85], Papanicolau-Varadhan [PV81]) in the elliptic case where the (tilted) Markov chain transition probabilities are assumed to be strictly positive (as in the case studied in [Y08]).

Upper bound. The upper bound part of the proof of the moment generating function starts with a "perturbationöf the exponential moment of the pair empirical measures $\mathfrak{L}_n$ defined in (2.1). This perturbation comes from integrating (in the exponential moment) certain "gradient functions"w.r.t. the local times $\mathfrak{L}_n$, while these gradient functions are intrinsically defined by the spatial action of the translation group $\mathbb{Z}^d$ on the environment space. In the elliptic case ([Y08], [R06] and [KRV06]), the class $\mathcal{K}$ of such gradient functions $F \in \mathcal{K}$ are required to satisfy the *closed* *loop condition* that underlines their gradient structure, a moment condition that requires $F \in L^{d+\varepsilon}(\mathbb{P})$ (which is related to the aforementioned moment assumption $\mathbb{E}^{\mathbb{P}}[|\log \pi|^{d+\varepsilon}] < \infty$) and a mean-zero condition that demands $\mathbb{E}^{\mathbb{P}}[F] = 0$. Exploiting these three properties, Rosenbluth ([R06]) proved that, for any $F \in \mathcal{K}$, the corresponding *correctors* $V_F(\omega, x) = \sum_{j=0}^{n-1} F(\tau_{x_j}\omega, x_{j+1} - x_j)$ defined as the integral of the gradient F along any path $x_0, \ldots, x_n = x$ between two fixed points x_0 and x_n (note that the choice of the path does not influence the integral, thanks to the aforementioned closed loop condition) have a "sub-linear growth at infinity". Roughly speaking, this means, $\mathbb{P}$-almost surely, $|V_F(\omega, x)| = o(|x|)$ as $|x| \to \infty$. This is a crucial technical step in Rosenbluth's work that is proved adapting the original approach of [KRV06] involving Sobolev embedding theorem and invoking Garsia-Rodemich-Rumsey estimate (the proof here hinges on the aforementioned moment condition $F \in L^{d+\varepsilon}(\mathbb{P})$). One then uses the ergodic theorem and take advantage of the mean-zero condition of the gradients F to get the desired sub-linearity property. This property implies, in particular, that the effect of the aforementioned perturbation caused by introducing the corrector in the exponential moment is indeed negligible. This is the crucial argument for the upper bound part.

Equivalence of lower and upper bounds. Having established both lower and upper bounds, one then faces the task of matching these two bounds. In the case of elliptic diffusions with a random drift, a seminal idea was introduced in [KRV06] by applying convex variational analysis followed by applications of certain min-max theorems. The success of this "min-maxäpproach relies on, among other requirements, "compactnessöf the underlying variational problem. In the elliptic case, this can be achieved by truncating the variational problem at a finite level which allows the application of the min-max theorems, followed by an approximation procedure by letting the truncation level to infinity. In the lattice, i.e., for elliptic RWRE a similar idea was used ([Y08], [R06]) in order to use the min-max argument. Indeed, by restricting the variational problem to a finite region in the environment space Ω and taking conditional expectation w.r.t. a finite σ-algebra B_k, [Y08] then used the min-max theorems for any fixed k. Roughly speaking, this leads to the study of conditional expectations

$$F_k := \mathbb{E}\big[f_k - f_k \circ \tau_e | B_{k-1}\big], \tag{2.7}$$

for test functions f_k, and one needs to prove that $F_k \to F$ as $k \to \infty$ such that $F \in \mathcal{K}$ (where $\mathcal{K}$ is the class of gradients with the required properties discussed in the upper bound part). Note that, for any fixed k, F_k is not a gradient. However, exploiting the underlying assumption $\mathbb{E}^{\mathbb{P}}[|\log \pi|^{d+\varepsilon}] < \infty$, one shows that $\{F_k\}_k$ remains uniformly bounded in $L^{d+\varepsilon}(\mathbb{P})$ so that one can take a weak limit F. After successive application of the tower property for the conditional expectations, one then proves that the limit F is indeed a gradient (i.e., satisfies the aforementioned closed loop condition), $F \in L^{d+\varepsilon}(\mathbb{P})$. Furthermore, $\mathbb{E}_{\mathbb{P}}[F] = 0$, which readily comes for free from (2.7) and the *invariant action* of τ_e w.r.t. the environment law $\mathbb{P}$. In particular, $F \in \mathcal{K}$ and modulo some technical work, this fact also matches the lower and upper bound of the limiting logarithmic moment generating function for the elliptic RWRE case.

Our approach for SRWPC. We now turn to a comparative description of the main strategy carried out for the proof of Theorem 2.3 for SRWPC in the present work, which, as a guiding philosophy, also follows the strategy of proving lower bound, upper bound and their equivalence.

Note that the lower bound follows the standard method of titling the environment Markov chain as the elliptic case. However, for the tilted environment Markov chain for the percolation models, the requisite ergodic theorem needs to be extended to the non-elliptic case which is the content of Theorem 3.2.

Now for the upper bound part for SRWPC, already the aforementioned moment condition of the elliptic case fails (zeroes of SRWPC transition probabilities π already make the first moment $\mathbb{E}_0(|\log \pi|)$ possibly infinite. Hence, we are not entitled to follow the method of Rosenbluth ([R06]) for proving the sub-linear growth property of the correctors. Moreover, the crucial mean-zero condition required in the elliptic case also fails for percolation due to the fundamental fact that the *spatial action of the shifts τ_e on Ω_0 is not $\mathbb{P}_0$-measure preserving.* The lack of these two properties necessarily forces us to reformulate the requisite conditions on our class of gradients. Besides the closed loop property in the infinite cluster, we demand uniform boundedness of the gradients in $\mathbb{P}_0$-norm and the validity of an "induced mean-zero property"to circumvent the above mentioned non-invariant nature of the spatial shifts τ_e w.r.t. $\mathbb{P}_0$, see Section 4.1 for details. With these assumptions, we prove the requisite ßub-linear growth"property of the correctors corresponding to our gradients, see Theorem 4.1. Note that our approach for proving this sub-linearity property is necessarily different from the one carried out by Rosenbluth ([R06]). For our purposes, we borrow techniques from ergodic theory, combined with geometric arguments that capture precise control of the “chemical distance"(the geodesic distance between two points x and y being in the infinite cluster $\mathcal{C}_\infty$ (proved in Lemma 4.2), as well as exponential tail bounds for the shortest distance between the origin and the first arrival of the cluster in the positive parts of the co-ordinate axes (proved in Lemma 4.3). Given the above sub-linear growth property on the infinite cluster which holds the pivotal argument, we then carry out the same “corrector perturbationäpproach as the elliptic case to the desired upper bound property, see Lemma 4.5.

Now for the “equivalence of bounds"for SRWPC, one can also try to emulate the strategy of ([Y08], [R06]) by carrying out the same convex variational analysis and applying the same min-max theorems by restricting to a finite region and conditional on a finite σ-algebra B_k. However, taking the conditional expectation as in (2.7) w.r.t. $\mathbb{E}_0$ any attempt towards deriving the requisite properties stated in Section 4.1 of the limiting function F completely fails. Note that in conditional expectation w.r.t. $\mathbb{E}_0$, one involves the measure $\mathbb{P}_0$ that is not preserved under the action of the shifts τ_e. In particular, we are not entitled to use any tower property. Plus, conditioning w.r.t. a finite σ-algebra B_k is incompatible for handling possibly long excursions of the infinite cluster before hitting the coordinate axes on each direction, which is a crucial issue one has to handle in order to prove the requisite induced mean-zero property of our limiting gradient.

Therefore, for the equivalence of bounds, we take a different route based on an *entropy coercivity* and *entropy penalization* method, which constitutes Section 5. This approach seems to be more natural in that it exploits the built-in structure of relative entropies that are already present in the underlying variational formulas. We make use of the coercivity property of the relative entropies in Lemma 5.2 and Lemma 5.3 to overcome the lack of the compactness in our variational analysis. One advantage of this method is that our variational analysis leads to the study of *gradients* directly, where we can work with functions

$$G_n(\omega, e) = g_n(\omega) - g_n(\tau_e\omega), \tag{2.8}$$

on the infinite cluster (see Lemma 5.4), instead of relying on conditional expectations like in (2.7).Given the gradient structure of G_n, and the estimates proved in Lemma 4.2 and 4.3, our analysis then also shows that the limiting gradients satisfy all the desired properties formulated in Section 4.1 (see Lemma 5.4) and the lower and upper bounds are readily matched. We also remark that the argument in our approach works equally well for the elliptic RWRE model considered before. Our method alleviates the effort needed in the earlier approach through the use of conditional expectations, tower property and Mazur's theorem in order to show that the limit of F_k defined in (2.7) is a gradient, and the equivalence of upper and lower bounds. In our approach, any weak limit of G_n defined in (2.8) is immediately a gradient and this readily makes the lower and the upper bound match (again, it is imperative here that we can work with G_n which is itself a gradient, unlike (2.7)). We refer to [R06, Sect.3.3] or [Y08, Sect.2.1.3] for a comparison with our approach in proving Theorem 5.1.

Remark 2 (Differences to the Kipnis-Varadhan corrector) Let us finally remark that the class of gradient functions introduced in Section 4.1 share some similarities to the gradient of Kipnis-Varadhan corrector which is a central object of interest for reversible random motions in random media. Particularly for SRWPC this is crucial for proving a quenched central limit theorem ([SS04], [MP07], [BB07], [PRS15])– the corrector expresses the deformation caused by a harmonic embedding of the random in the infinite cluster in $\mathbb{R}^d$, and modulo this deformation, the random walk becomes a martingale. However, our gradient functions that are defined in Section 4.1 are structurally different from the gradient of the Kipnis-Varadhan corrector. Though they share similar properties as *gradients*, our gradients miss the above mentioned *harmonicity* property enjoyed by the Kipnis-Varadhan corrector. This can be explained by the fact that large deviation lower bounds are based on a certain *tilt* which spoils any inherent reversibility of the model, which is a crucial base of Kipnis-Varadhan theory. □

The rest of the article is organized as follows. In Section 3, Section 4 and Section 5 we prove the lower bound, the upper bound and the equivalence of bounds for Theorem 2.3, respectively. Section 6 is devoted to the proofs of Theorem 2.3, Theorem 2.1, Corollary 2.4 and Lemma 2.2.

3. Lower bounds of Theorem 2.1 and Theorem 2.3

We first introduce a class of environment Markov chains for SRWPC and prove an ergodic theorem for these in Section 3.1. We then derived the lower bounds for Theorem 2.1 and Theorem 2.3 in Section 3.2.

3.1 An ergodic theorem for Markov chains on non-elliptic environments

In this section we need some input from the environment seen from the particle, which, with respect to a suitably changed measure, possesses important ergodic properties.

Recall that, given the transition probabilities π from (1.4), for $\mathbb{P}_0$- almost every $\omega \in \Omega_0$, the process $(\tau_{X_n}\omega)_{n\geq 0}$ is a Markov chain with transition kernel

$$(R_\pi g)(\omega) = \sum_{e\in\mathbb{B}_d} \pi_\omega(0,e)g(\tau_e\omega),$$

for every function g on Ω_0 which is measurable and bounded.

We need to introduce a class of transition kernels on the space of environments. We denote by $\widetilde{\Pi}$ the space of functions $\tilde{\pi} : \Omega_0 \times \mathbb{B}_d \to [0,1]$ which are measurable in Ω_0, $\sum_{e\in\mathbb{B}_d} \tilde{\pi}(\omega,e) = 1$ for almost every $\omega \in \Omega_0$ and for any $\omega \in \Omega_0$ and $e \in \mathbb{B}_d$,

$$\tilde{\pi}(\omega,e) = 0 \text{ if and only if } \pi_\omega(0,e) = 0. \tag{3.1}$$

For any $\tilde{\pi} \in \widetilde{\Pi}$ and $\omega \in \Omega_0$, we define the corresponding quenched probability distribution of the Markov chain $(X_n)_{n\geq 0}$ by

$$\begin{aligned} &P_0^{\tilde{\pi},\omega}(X_0 = 0) = 1 \\ &P_0^{\tilde{\pi},\omega}(X_{n+1} = x+e | X_n = x) = \tilde{\pi}(\tau_x\omega, e). \end{aligned} \tag{3.2}$$

With respect to any $\tilde{\pi} \in \widetilde{\Pi}$ we also have a transitional kernel

$$(R_{\tilde{\pi}}g)(\omega) = \sum_{e\in\mathbb{B}_d} \tilde{\pi}(\omega,e)g(\tau_e\omega),$$

for every measurable and bounded g. For any measurable function $\phi \geq 0$ with $\int \phi \mathrm{d}\mathbb{P}_0 = 1$, we say that the measure $\phi\mathrm{d}\mathbb{P}_0$ is $R_{\tilde{\pi}}$-invariant, or simply $\tilde{\pi}$-invariant, if,

$$\phi(\omega) = \sum_{e\in\mathbb{B}_d} \tilde{\pi}\big(\tau_{-e}\omega, e\big)\phi\big(\tau_{-e}\omega\big). \tag{3.3}$$

Note that in this case,

$$\int g(\omega)\phi(\omega)d\mathbb{P}_0(\omega) = \int (R_{\tilde{\pi}}g)(\omega)\phi(\omega)d\mathbb{P}_0(\omega), \tag{3.4}$$

for every bounded and measurable g.

We denote by $\mathcal{E}$ such pairs of $(\tilde{\pi},\phi)$, i.e.,

$$\mathcal{E} = \bigg\{(\tilde{\pi},\phi) \colon\ \tilde{\pi} \in \widetilde{\Pi}, \phi \geq 0, \mathbb{E}_0(\phi) = 1,\ \phi\mathrm{d}\mathbb{P}_0 \,\mathrm{is}\ \tilde{\pi} -\ \mathrm{invariant}\bigg\}. \tag{3.5}$$

We need an elementary lemma which we will be using frequently. Recall the set $\mathcal{M}_1^\star$ from (2.3).

Lemma 3.1. *There is a one-to-one correspondence between the sets $\mathcal{M}_1^\star$ and $\mathcal{E}$.*

Proof. Given any $(\tilde{\pi},\phi) \in \mathcal{E}$, we take

$$\mathrm{d}\mu(\omega,e) = \tilde{\pi}(\omega,e)\phi(\omega)\,\mathrm{d}\mathbb{P}_0 = \tilde{\pi}(\omega,e)\bigg(\sum_{e:\tau_e\omega'=\omega} \tilde{\pi}(\omega',e)\phi(\omega')\bigg)\,\mathrm{d}\mathbb{P}_0. \tag{3.6}$$

By (2.2), $\mathbb{P}_0$-almost surely,

$$\mathrm{d}(\mu)_1(\omega) = \sum_{e\in\mathbb{B}_d} \mathrm{d}\mu(\omega, e) = \sum_{e:\tau_e\omega'=\omega} \tilde{\pi}(\omega', e)\phi(\omega')\,\mathrm{d}\mathbb{P}_0 = \sum_{e:\tau_e\omega'=\omega} \mathrm{d}\mu(\omega', e) = \mathrm{d}(\mu)_2(\omega).$$

Hence, $(\mu)_1 = (\mu)_2 \ll \mathbb{P}_0$. Furthermore, if the edge $0 \leftrightarrow e$ is present in the configuration ω (i.e., $\omega(e) = 1$), then $\pi_\omega(0, e) > 0$, and by our requirement (3.1),

$$\frac{\mathrm{d}\mu(\omega, e)}{\mathrm{d}(\mu_1)(\omega)} = \tilde{\pi}(\omega, e) > 0,$$

Hence $\mu \in \mathcal{M}_1^\star$. Conversely, given any $\mu \in \mathcal{M}_1^\star$, we can choose $(\tilde{\pi}, \phi) = \big(\frac{\mathrm{d}\mu}{\mathrm{d}(\mu)_1}, \frac{\mathrm{d}(\mu)_1}{\mathrm{d}\mathbb{P}_0}\big)$ and readily check that $(\tilde{\pi}, \phi) \in \mathcal{E}$. □

We now state and prove the following ergodic theorem for the environment Markov chain under any transition kernel $\tilde{\pi} \in \widetilde{\Pi}$. Theorem 3.2 is an extension of a similar statement (see Kozlov [K85], Papanicolau-Varadhan [PV81]) that holds for elliptic transition kernels $\widetilde{\pi}(\cdot, e)$ to the non-elliptic case.

Theorem 3.2. *Fix $\tilde{\pi} \in \widetilde{\Pi}$. If there exists a probability measure $\mathbb{Q} \ll \mathbb{P}_0$ which is $\tilde{\pi}$-invariant, then $\mathbb{Q} \sim \mathbb{P}_0$ and the environment Markov chain with initial law $\mathbb{Q}$ and transition kernel $\tilde{\pi}$ is stationary and ergodic for $\mathbb{P}_0$. Moreover, there is at most one probability measure $\mathbb{Q}$ which is $\tilde{\pi}$- invariant probability and is absolutely continuous with respect to $\mathbb{P}_0$.*

Proof. We fix $\tilde{\pi} \in \widetilde{\Pi}$ and let $\mathbb{Q} \ll \mathbb{P}_0$ be $\tilde{\pi}$- invariant. We prove the theorem in three steps.

Let us first show that, $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}_0} > 0$ $\mathbb{P}_0$- almost surely. This will imply that $\mathbb{Q} \sim \mathbb{P}_0$.

Indeed, to the contrary, let us assume that, $0 < \mathbb{P}_0(A) < 1$ where $A = \big\{\omega\colon \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}_0}(\omega) > 0\big\}$. Then, $\mathbb{Q} \sim \mathbb{P}_0(\cdot|A)$. If we sample $\omega_1 \in \Omega_0$ according to $\mathbb{Q}$ and ω_2 according to $\tilde{\pi}(\omega_1, \cdot)$, then the distribution of ω_2 is absolutely continuous with respect to $\mathbb{Q}$ (recall $\mathbb{Q}$ is $\tilde{\pi}$ invariant) and thus, on A^c, the distribution of ω_2 has zero measure.

This implies that, for almost every $\omega_1 \in A$ and every $e \in \mathbb{B}_d$ such that $\tilde{\pi}(\omega_1, e) > 0$, $\tau_e\omega_1 \in A$. Since $\tilde{\pi} \in \widetilde{\Pi}$, for almost every $\omega_1 \in A$ and every $e \in \mathbb{B}_d$ such that $\pi(\omega_1, e) > 0$, $\tau_e\omega_1 \in A$. Now if we sample ω_1 according to $\mathbb{P}_0(\cdot|A)$ and ω_2 according to $\pi(\omega_1, \cdot)$, then, with probability 1, $\omega_2 \in A$. In other words, A is invariant under π (more precisely, A is invariant under the Markov kernel R_π). Since $\mathbb{P}_0$ is π-ergodic (see [BB07, Proposition 3.5]), $\mathbb{P}_0(A) \in \{0, 1\}$. By our assumption, $\mathbb{P}_0(A) = 1$.

Now we prove that the environment Markov chain with initial law $\mathbb{Q}$ and transition kernel $\tilde{\pi}$ is $\mathbb{P}_0$ ergodic. Let us assume on the contrary, that for some measurable D, $\mathbb{Q}(D) > 0$, $\mathbb{Q}(D^c) > 0$ and D is $\tilde{\pi}$ invariant. Hence $\mathbb{P}_0(D) > 0$ and $\mathbb{P}_0(D^c) > 0$, by $\mathbb{Q} \sim \mathbb{P}_0$.Further, the conditional measure $\mathbb{Q}_D(\cdot) = \mathbb{Q}(\cdot|D)$ is $\tilde{\pi}$ invariant and $\mathbb{Q}_D \ll \mathbb{P}_0$. But $\mathbb{Q}_D(D^c) = 0$ and hence, $\frac{\mathrm{d}\mathbb{Q}_D}{\mathrm{d}\mathbb{P}_0}(D^c) = 0$. This contradicts the first step.

To conclude the proof, we need to prove uniqueness of any $\mathbb{Q}$ which is $\tilde{\pi}$- invariant and absolutely continuous with respect to $\mathbb{P}_0$. Let $\Omega^{\mathbb{Z}}$ be the space of the trajectories $(\dots, \omega_{-1}, \omega_0, \omega_1, \dots)$ of the environment chain, $\mu_{\mathbb{Q}}$ the measure associated to the transition kernel $\tilde{\pi}$ whose finite dimensional distributions are given by

$$\mu_{\mathbb{Q}}\big((\omega_{-n}, \dots, \omega_n) \in A\big) = \int_A \mathbb{Q}(\mathrm{d}\omega_{-n}) \prod_{j=-n}^{n-1} \tilde{\pi}\big(\omega_j, \mathrm{d}\omega_{j+1}\big).$$

for any finite dimensional cylinder set A in $\Omega^{\mathbb{Z}}$. Let $T : \Omega^{\mathbb{Z}} \longrightarrow \Omega^{\mathbb{Z}}$ be the shift given by $(T\omega)_n = \omega_{n+1}$ for all $n \in \mathbb{Z}$. Since $\mathbb{Q}$ is $\tilde{\pi}$- invariant and ergodic, by Birkhoff's theorem,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} g \circ T^k = \int g \mathrm{d}\mu_{\mathbb{Q}},$$

$\mu_{\mathbb{Q}}$ (and hence $\mu_{\mathbb{P}_0}$) almost surely for any bounded and measurable g on $\Omega^{\mathbb{Z}}$. Since the environment chain $(\tau_{X_k}\omega)_{k\geq 0}$ has the same law in $\int P_0^{\tilde{\pi},\omega}\mathrm{d}\mathbb{Q}$ as $(\omega_0,\omega_1,\dots)$ has in $\mu_{\mathbb{Q}}$, if $f(\omega_0)=g(\omega_0,\omega_1,\dots)$, then

$$\lim_{n\to\infty}\sum_{k=0}^{n-1} f\circ\tau_{X_k}=\lim_{n\to\infty}\sum_{k=0}^{n-1} g\circ T^k=\int g\mathrm{d}\mu_{\mathbb{Q}}=\int f\mathrm{d}\mathbb{Q},$$

for any bounded and measurable f on Ω. The uniqueness of $\mathbb{Q}$ follows. □

Corollary 3.3. *For any pair $(\tilde{\pi},\phi)\in\mathcal{E}$ and every continuous and bounded function $f:\Omega_0\times\mathbb{B}_d\to\mathbb{R}$,*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(\tau_{X_k}\omega,X_{k+1}-X_k)=\int_{\Omega_0}\mathrm{d}\mathbb{P}_0\,\phi(\omega)\sum_e f(\omega,e)\tilde{\pi}(\omega,e),\ \ \mathbb{P}_0\times P_0^{\tilde{\pi},\omega}\text{-a.s.}$$

Proof. This is an immediate consequence of Theorem 3.2 and Birkhoff's ergodic theorem. □

3.2 Proof of lower bounds. We now prove the required lower bound (2.6). Its proof follows a standard change of measure argument and given Theorem 3.2, although the argument is very similar to Yilmaz ([Y08]), we present this short proof for convenience of the reader and to keep the article self-contained. Recall the definition of $\mathfrak{I}$ from (2.4).

Lemma 3.4 (The lower bound)**.** *For any open set $\mathcal{G}$ in $\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$, $\mathbb{P}_0$- almost surely,*

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log P_0^{\pi,\omega}\big(\mathfrak{L}_n\in\mathcal{G}\big)&\geq-\inf_{\mu\in\mathcal{G}}\mathfrak{I}(\mu)\\&=-\inf_{\mu\in\mathcal{G}}\mathfrak{I}^{\star\star}(\mu).\end{aligned}\tag{3.7}$$

Proof. For the lower bound in (3.7), it is enough to show that, for any $\mu\in\mathcal{M}_1^\star$ and any open neighborhood $\mathcal{U}$ containing μ,

$$\liminf_{n\to\infty}\frac{1}{n}\log P_0^{\pi,\omega}\big(\mathfrak{L}_n\in\mathcal{U}\big)\geq-\mathfrak{I}(\mu).\tag{3.8}$$

Given $\mu\in\mathcal{M}_1^\star$, from Lemma 3.1 we can get the pair

$$(\tilde{\pi},\phi)=\left(\frac{\mathrm{d}\mu}{\mathrm{d}(\mu)_1},\frac{\mathrm{d}(\mu)_1}{\mathrm{d}\mathbb{P}_0}\right)\in\mathcal{E},\tag{3.9}$$

and by Theorem 3.2,

$$\lim_{n\to\infty}P_0^{\tilde{\pi},\omega}\big(\mathfrak{L}_n\in\mathcal{U}\big)=1.\tag{3.10}$$

Further,

$$\begin{aligned}P_0^{\pi,\omega}\big(\mathfrak{L}_n\in\mathcal{U}\big)&=E_0^{\tilde{\pi},\omega}\left\{1\!\!1_{\{\mathfrak{L}_n\in\mathcal{U}\}}\frac{\mathrm{d}P_0^{\pi,\omega}}{\mathrm{d}P_0^{\tilde{\pi},\omega}}\right\}\\&=\int\mathrm{d}P_0^{\tilde{\pi},\omega}\left\{1\!\!1_{\{\mathfrak{L}_n\in\mathcal{U}\}}\exp\left\{-\log\frac{\mathrm{d}P_0^{\tilde{\pi},\omega}}{\mathrm{d}P_0^{\pi,\omega}}\right\}\right\}.\end{aligned}$$

Hence, by Jensen's inequality,

$$\begin{aligned}\liminf_{n\to\infty}\frac{1}{n}\log P_0^{\pi,\omega}\big(\mathfrak{L}_n\in\mathcal{U}\big)&\geq\liminf_{n\to\infty}\frac{1}{n}\log P_0^{\tilde{\pi},\omega}\big(\mathfrak{L}_n\in\mathcal{U}\big)\\&\qquad-\limsup_{n\to\infty}\frac{1}{nP_0^{\tilde{\pi},\omega}\big(\mathfrak{L}_n\in\mathcal{U}\big)}\int_{\{\mathfrak{L}_n\in\mathcal{U}\}}\mathrm{d}P_0^{\tilde{\pi},\omega}\left\{\log\frac{\mathrm{d}P_0^{\tilde{\pi},\omega}}{\mathrm{d}P_0^{\pi,\omega}}\right\}\\&=-\int\mathrm{d}\mathbb{P}_0(\omega)\,\phi(\omega)\sum_{|e|=1}\tilde{\pi}(\omega,e)\log\frac{\tilde{\pi}(\omega,e)}{\pi_\omega(0,e)}\\&=-\mathfrak{I}(\mu),\end{aligned}$$

where the first equality follows from (3.10) and corollary 3.3 and the second equality follows from (3.9). This proves (3.8). Finally, since $\mathcal{G}$ is open, $\inf_{\mu\in\mathcal{G}}\mathfrak{I}(\mu) = \inf_{\mu\in\mathcal{G}}\mathfrak{I}^{\star\star}(\mu)$ (see [R70]). This proves the equality in (3.7) and the lemma. $\square$

We now prove the lower bound for the limiting logarithmic moment generating function required for Theorem 2.3.

Corollary 3.5. *For every continuous and bounded function $f : \Omega_0 \times \mathbb{B}_d \longrightarrow \mathbb{R}$ and for $\mathbb{P}_0$-almost every $\omega \in \Omega_0$,*

$$\liminf_{n\to\infty} \frac{1}{n} \log E_0^{\pi,\omega}\bigg\{ \exp\bigg(\sum_{k=0}^{n-1} f\big(\tau_{X_k}\omega, X_{k+1}-X_k\big)\bigg)\bigg\} \ge \sup_{\mu\in\mathcal{M}_1^\star} \{\langle f,\mu\rangle - \mathfrak{I}(\mu)\} = \sup_{\mu\in\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)} \big\{\langle f,\mu\rangle - \mathfrak{I}(\mu)\big\}. \tag{3.11}$$

Proof. This follows immediately from Varadhan's lemma and Lemma 3.4. $\square$

We denote the variational formula in Corollary 3.5 by

$$\begin{aligned} \overline{H}(f) &= \sup_{\mu\in\mathcal{M}_1^\star} \{\langle f,\mu\rangle - \mathfrak{I}(\mu)\} \\ &= \sup_{(\tilde\pi,\phi)\in\mathcal{E}} \bigg\{ \int \mathrm{d}\mathbb{P}_0(\omega)\,\phi(\omega) \sum_{|e|=1} \tilde\pi(\omega,e)\bigg\{ f(\omega,e) - \log\frac{\tilde\pi(\omega,e)}{\pi_\omega(0,e)}\bigg\}\bigg\}, \end{aligned} \tag{3.12}$$

recall from Lemma 3.1 the one-to-one correspondence between elements of the set $\mathcal{M}_1^\star$ and the pairs $\mathcal{E}$ (and (3.9), (2.4)). For the variational analysis that follows in Section 5, it is convenient to write down a more tractable representation of the above variational formula. This is based on the following observation, which was already made by Kosygina-Rezakhanlou-Varadhan ([KRV06]) and used ivy Yilmaz ([Y08]) and Rosenbluth ([R06]). Recall that by (3.4), if $(\phi,\tilde\pi)\in\mathcal{E}$, then for any bounded and measurable function g on Ω_0,

$$\sum_e \int \phi(\omega)\tilde\pi(\omega,e)\big(g(\omega) - g(\tau_e\omega)\big)\mathrm{d}\mathbb{P}_0(\omega) = 0. \tag{3.13}$$

On the other hand, if $(\phi,\tilde\pi)\notin\mathcal{E}$, then for some bounded and measurable function g on Ω_0, the above integral on the left hand side is non-zero. By taking constant multiples of such a function g, then

$$\inf_g \int \phi(\omega) \sum_{e\in\mathbb{B}_d} \tilde\pi(\omega,e)\big(g(\omega) - g(\tau_e\omega)\big)\,\mathrm{d}\mathbb{P}_0(\omega) = \begin{cases} 0 & \text{if } (\phi,\tilde\pi)\in\mathcal{E} \\ -\infty & \text{else.} \end{cases}$$

with the infimum being taken over any bound and measurable function g. Hence, we can rewrite (3.12) as

$$\overline{H}(f) = \sup_\phi \sup_{\tilde\pi\in\widetilde\Pi} \inf_g \bigg[\int \mathrm{d}\mathbb{P}_0(\omega)\,\phi(\omega)\bigg\{ \sum_{e\in\mathbb{B}_d} \tilde\pi(\omega,e)\bigg(f(\omega,e) - \log\frac{\widetilde\pi_\omega(0,e)}{\pi_\omega(0,e)} + \big(g(\omega) - g(\tau_e\omega)\big)\bigg)\bigg\}\bigg] \tag{3.14}$$

4. Upper bound for the proof of Theorem 2.3

We will now introduce the class of relevant gradient functions in Section 4.1, derive an important property of these functions in Section 4.2 and prove the desired upper bound of Theorem 2.3 in Section 4.3.

4.1 The class $\mathcal{G}_\infty$ of gradients and the corresponding correctors

We introduce a class of functions which will play an important role for the large deviation analysis to follow. However, before introducing this class we need the notion of the *induced shift* on Ω_0.

Fix $\omega \in \Omega_0$ and $e \in \mathbb{B}_d$. Let

$$k(\omega, e) = \inf\{k \geq 1 : \tau_{ke}\, \omega \in \Omega_0\}. \tag{4.1}$$

We would like to argue that $k(\omega, e)$ is finite $\mathbb{P}_0$-almost surely. This is immediate from the mixing property (i.e., the property which demands $\big|\mathbb{P}[x, y \in \mathcal{C}_\infty] - \mathbb{P}[x \in \mathcal{C}_\infty]\mathbb{P}[y \in C_\infty]\big| \to 0$ as $|x - y| \to \infty$) present in all the percolation models (recall (1.3)). Hence by the spatial erodicity along any of the coordinate directions and by Birkhoff's ergodic theorem, for each $e \in \mathbb{B}_d$, the set $\{k \geq 1 : \tau_{ke}\, \omega \in \Omega_0\}$ has positive density in $\mathbb{N}$. Hence, $k(\omega, e)$ is finite $\mathbb{P}_0$-almost surely.

Then the *induced shift* is defined as

$$\sigma_e(\omega) = \tau_{k(\omega,e)e}\, \omega. \tag{4.2}$$

It is well-known that, for every $e \in \mathbb{B}_d$, $\sigma_e \colon \Omega_0 \to \Omega_0$ is $\mathbb{P}_0$-measure preserving and ergodic. Furthermore, for any $k \in \mathbb{N}$, we inductively set

$$n_1(\omega, n) = k(\omega, e) \qquad n_{k+1}(\omega, e) = n_k(\sigma_e\omega, e). \tag{4.3}$$

Now we turn to the definition of $\mathcal{G}_\infty$. We say that a function $G : \Omega_0 \times \mathbb{B}_d \longrightarrow \mathbb{R}$ is in class $\mathcal{G}_\infty$ if it satisfies the conditions (4.4), (4.5) and (4.10) listed below:

- **Uniform boundedness.** For every $e \in \mathbb{B}_d$,

$$\operatorname*{ess\,sup}_{\mathbb{P}_0} G(\cdot, e) = A < \infty. \tag{4.4}$$

- **Closed loop on the cluster.** Let $(x_0, \dots, x_n)$ be a closed loop on the infinite cluster $\mathcal{C}_\infty$ (i.e., $x_0, x_1, \dots, x_n$ is a nearest neighbor occupied path so that $x_0 = x_n$). Then,

$$\sum_{j=0}^{n-1} G(\tau_{x_j}\omega, x_{j+1} - x_j) = 0 \quad \mathbb{P}_0 - \text{almost surely.} \tag{4.5}$$

  For any $G \in \mathcal{G}_\infty$, the closed loop condition has two important consequences. First, along any nearest neighbor occupied path $(x_0, x_1, \dots, x_n)$ so that $x_0 = 0$ and $x_n = x$ on $\mathcal{C}_\infty$, for any $G(\cdot, \cdot)$ that satisfies (4.5), we can define the *corrector* corresponding to G as

$$V(\omega, x) = V_G(\omega, x) = \sum_{j=0}^{n-1} G(\tau_{x_j}\omega, x_{j+1} - x_j). \tag{4.6}$$

  By (4.5), this definition is clearly independent of the chosen path for almost every $\omega \in \{x \in \mathcal{C}_\infty\}$. Also note that, for any G that satisfies (4.5), $V = V_G$ satisfies the following *Shift covariance* condition: For $\mathbb{P}_0$-almost every $\omega \in \Omega_0$ and all $x, y \in \mathcal{C}_\infty$,

$$V(\omega, x) - V(\omega, y) = V(\tau_y\omega, x - y). \tag{4.7}$$

- **Zero induced mean:** Recall the definition of $k(\omega, e)$ from (4.1) and write

$$v_e = k(\omega, e)\, e \tag{4.8}$$

  for every $\omega \in \Omega$ and $e \in \mathbb{B}_d$. Let $\big\{0 = x_0, x_1, \dots, x_k = k(\omega, e)\, e\big\}$ be an ω-open path from 0 to $k(\omega, e)\, e$. For any $G(\cdot, \cdot)$ that satisfies (4.5), we again write

$$V(\omega, v_e) = V_G(\omega, v_e) = \sum_{i=0}^{k-1} G(\tau_{x_i}\omega, x_{i+1} - x_i). \tag{4.9}$$

  Again, the choice of the path doesn't influence $V(\omega, e)$. We then say that $V = V_G$ satisfies the *induced zero mean property* by requiring that for every $e \in \mathbb{B}_d$,

$$\mathbb{E}_0\big[V(\cdot, e)\big] = 0. \tag{4.10}$$

4.2 Sub-linear growth of the correctors at infinity

This section is devoted to the proof of the following important property of functions in the class $\mathcal{G}_\infty$.

Theorem 4.1 (Sub-linear growth at infinity on the cluster)**.** *For any $G \in \mathcal{G}_\infty$, $V = V_G$ has at most sub-linear growth at infinity on the infinite cluster $\mathbb{P}_0$- almost surely,. In other words,*

$$\lim_{n\to\infty} \max_{\substack{x\in\mathcal{C}_\infty \\ |x|\le n}} \frac{|V(\omega,x)|}{n} = 0.$$

Before we present the proof of Theorem 4.1, which is carried out at the end of this section, we need some important estimates related to the geometry of the infinite percolation cluster $\mathcal{C}_\infty$ presented in the following two lemmas. Lemma 4.2 gives a precise bound on the shortest distance of two points in the infinite cluster (the *chemical distance*) and Lemma 4.3 gives an exponential tail bound on the graph distance between the origin and the the first arrival $v_e = k(\omega, e)e$ (recall (4.8)), of the cluster on the positive part of any of the coordinate directions. Both lemmas are well-known in the literature covering i.i.d. Bernoulli percolation and the site percolation model discussed in Section 1.2. For the random cluster model, contents of these two results are new, to the best of our knowledge. Apart from the proof of Theorem 4.1, both lemmas will be helpful in carrying out our variational analysis in Section 5 (see the proof of Lemma 5.6).

We first turn to the following estimate on the *chemical distance* $\mathrm{d_{ch}}(x,y) = \mathrm{d_{ch}}(\omega;\, x,y)$ of two points $x,y \in \mathcal{C}_\infty$, which is defined to be the minimal length of an ω-open path connecting x and y in the configuration $\omega \in \Omega_0$. The following result, originally proved by Antal and Pisztora ([AP96]) for supercritical i.i.d. Bernoulli percolation, asserts that the chemical distance of two points in the cluster is comparable to their Euclidean distance.

Lemma 4.2. *Fix $\delta > 0$. Then there exists a constant $\rho = \rho(p,d)$ such that, $\mathbb{P}_0$- almost surely, for every n large enough and points $x, y \in \mathcal{C}_\infty$ with $|x| < n, |y| < n$ and $|x-y| < \delta n$, we have $\mathrm{d_{ch}}(x,y) < \rho\delta n$.*

Proof. Let us first treat the bond percolation models. For i.i.d. Bernoulli percolation model (recall Section 1.1), the statement of this lemma follows from the classical estimate of Antal-Pisztora (Theorem 1.1, [AP96])).

We now prove the lemma for the supercritical random-cluster model. Recall that we assume that $p > \widehat{p}_c(q)$. For any $r \ge 0$, we fix a box $B_0(r) := [-r,r]^d$ and set, for any $z \in \mathbb{Z}^d$,

$$B_z(N) = \tau_{(2N+1)z} B_0(N) \qquad B'_z(N) := \tau_{(2N+1)z} B_0(5N/4).$$

Here τ_z is the transformation on $\mathbb{Z}^d$ defined by $\tau_z(x) = z + x$. We define $R_i^{(N)}$ to be the event in $\{0,1\}^{\mathbb{B}_d}$ satisfying the following three conditions:

- There exists a *unique crossing open cluster* for $B'_z(N)$. In other words, there is a connected subset $\mathcal{C}$ of an open cluster such that it is contained in $B'_z(N)$, and, for all d directions there is a path in $\mathcal{C}$ connecting the left face and the right face of $B'_z(N)$.
- The cluster in the above requirement intersects all boxes with diameter larger than $N/10$.
- All open clusters with diameter larger than $N/10$ are connected in $B'_z(N)$.

Recall the measures $\mathbb{P}^{(\xi)}_{\Lambda,p,q}$ and $\mathbb{P}^{(b)}_{p,q}$ corresponding to the random cluster model. Then, under the map

$$\phi_N : \{0,1\}^{\mathbb{B}_d} \to \{0,1\}^{\mathbb{Z}^d} \qquad (\phi_N\omega)_z = \mathbb{1}_{R_z^{(N)}}(\omega) \quad \forall\, z \in \mathbb{Z}^d,$$

we let

$$\mathbb{P}^{(b)}_{p,q,N} = \mathbb{P}^{(b)}_{p,q} \circ \phi_N^{-1}$$

to be the image measure of $\mathbb{P}^{(b)}_{p,q}$. By [P96, Theorem 3.1] for $d \ge 3$ and [CM04, Theorem 9] for $d = 2$, we see that there exist constants $c'_1, c'_2 > 0$ (depending only on d, p and q), such that for any $N \ge 1$ and $i \in \mathbb{Z}^d$,

$$\sup_{\xi\in\Omega} \mathbb{P}^{(\xi)}_{B'_z(N),p,q}\left[\left(R_i^{(N)}\right)^c\right] \le c'_1\, \mathrm{e}^{-c'_2 N}.$$

Let $Y_z : \{0,1\}^{\mathbb{Z}^d} \to \{0,1\}$ be the projection mapping to the coordinate $z \in \mathbb{Z}^d$. By using the DLR property for the random-cluster model ([G06, Section 4.4]),

$$\lim_{N\to\infty} \sup_{z\in\mathbb{Z}^d} \text{ess.sup}\ \mathbb{P}^{(b)}_{p,q,N} \left[Y_z = 0 \middle| \sigma\bigg(Y_x : |x-z|_\infty \geq 2\bigg)\right] = 0.$$

By using [LSS97, Theorem 1.3], we see that there exists a function $\overline{p}(\cdot) : \mathbb{N} \to [0,1)$ such that $\overline{p}(N) \to 1$ as $N \to \infty$ and the Bernoulli product measure

$$\mathbb{P}^\star_{\overline{p}(N)} = \big(\overline{p}(N)\ \delta_1\ +\ \big(1-\overline{p}(N)\big)\delta_0\big)^{\mathbb{Z}^d}$$

on $\{0,1\}^{\mathbb{Z}^d}$ with parameter $\overline{p}(N)$ is dominated by $\mathbb{P}^{(b)}_{p,q,N}$ for each N, i.e., for any increasing event A,

$$\mathbb{P}^\star_{\overline{p}(N)}(A) \leq \mathbb{P}^{(b)}_{p,q,N}(A).$$

The above estimate now enable us to deduce that, for some positive integer $k \leq |x-y|$, and some constant $c = c(p,q,d) > 0$ there exist macroscopic clusters $C^\star_i$ (see p. 1047, [AP96])

$$\begin{aligned}\mathbb{P}_0\bigg\{\mathrm{d_{ch}}(x,y) > \rho|x-y|,\ x,y \in \mathcal{C}_\infty\bigg\} &\leq \mathbb{P}^{(b)}_{p,q,N}\bigg\{\frac{1}{k+1}\sum_{i=0}^k \big(|C^\star_i|+1\big) > \rho\,|x-y|\ \frac{c\,N^{-d}}{k+1}\bigg\} \\ &\leq \mathbb{P}^\star_{\overline{p}(N)}\bigg\{\frac{1}{k+1}\sum_{i=0}^k \big(|C^\star_i|+1\big) > \rho\,|x-y|\ \frac{c\,N^{-d}}{k+1}\bigg\} \\ &\leq \mathbb{P}^\star_{\overline{p}(N)}\bigg\{\frac{1}{k+1}\sum_{i=0}^k \big(\widetilde{C}^\star_i+1\big) > \rho\,|x-y|\ \frac{c\,N^{-d}}{k+1}\bigg\}\end{aligned} \tag{4.11}$$

for some i.i.d. random subsets $\{\widetilde{C}^\star_i\}_i$ in $\mathbb{B}_d$ (called "preclustersïn [AP96]) with the property that

$$\mathbb{E}^\star_{\overline{p}(N)}\left[\mathrm{e}^{\lambda\big[|C^\star_i|+1\big]}\right] < \infty \qquad \text{for any } i \ \text{ and some } \ \lambda > 0.$$

We now apply Chebyshev's inequality to the last probability in (4.11) and choose ρ suitably so that

$$\mathbb{P}_0\bigg\{\mathrm{d_{ch}}(x,y) > \rho|x-y|,\ x,y \in \mathcal{C}_\infty\bigg\} \leq \mathrm{e}^{-c|x-y|}. \tag{4.12}$$

By Borel-Cantelli lemma, we conclude the proof of Lemma 4.2 for random cluster models.

For the site percolation models (i.e., random interlacements, its vacant set and the level sets of the Gaussian free field) introduced in Section 1.2, Lemma 4.2 follows from the estimate

$$\mathbb{P}_0\bigg\{\mathrm{d_{ch}}(x,y) \geq \rho\,|x-y|,\ x,y \in \mathcal{C}_\infty\bigg\}\ \leq c_1\ \mathrm{e}^{-c_1\ (\log|x-y|)^{1+c_2}},$$

for constants $c_1, c_2 > 0$ and any $x \in \mathbb{Z}^d$. This statement and its proof can be found in [DRS14, Theorem 1.3]. This concludes the proof of Lemma 4.2. □

For any $e \in \mathbb{B}_d$, from (4.8) we recall that $v_e = k(\omega, e)e$. Let $\ell = \ell(\omega)$ denote the shortest path distance from 0 to v_e. Then we have the following tail estimate on ℓ:

Lemma 4.3. *For some constant $c_1, c_2 > 0$,*

$$\mathbb{P}_0\big\{\ell > n\big\} \leq c_1\,\mathrm{e}^{-c_2 n}.$$

Proof. Lemma 4.3 follows from ([BB07, Lemma 4.3]) for i.i.d. Bernoulli bond and site percolations, and from [PRS15, Section 5] for the site percolation models appearing in Section 1.2.

We turn to the requisite estimate corresponding to the random cluster model defined in Section 1.1. Let us first handle the case $d \geq 3$ and recall the definition of slab-critical probability $\widehat{p}_c(q)$ from (1.1) and recall that we assume $p > \widehat{p}_c(q)$. Then we can take a large number L so that $p > \widehat{p}_c(q, L)$ and $[0, L-1] \times \mathbb{Z}^{d-1}$ contains an infinite cluster, which is a subset of the unique infinite cluster $\mathcal{C}_\infty$. For any $e \in \mathbb{B}_d$, we recall the definition of $k(\omega, e)$ from (4.1) and note that we write $v_e = k(\omega, e)e$. Then,

$$\Big\{|v_e| \geq Ln;\, 0 \in \mathcal{C}_\infty\Big\} \subset \bigcap_{i=1}^n \tau_{iLe}(A_L^c), \tag{4.13}$$

where

$$A_L := \bigcup_{j=1}^L \Big\{0 \leftrightarrow je \text{ in } [0, L-1] \times \mathbb{Z}^{d-1}\Big\}.$$

We also define, for $m \geq 1$,

$$A_{L,m} := \bigcup_{j=1}^L \{0 \leftrightarrow je \text{ in } S(L,m)\}.$$

Then $\{A_L\}_{L>0}$ and $\{A_{L,m}\}_{m\geq 1}$ are increasing events. Furthermore, we note that the random cluster model satisfies the *finite-energy property*. This means, for any $\delta \in (0, 1/2)$, the conditional probability for an edge to be open, knowing the states of all the other edges, is bounded away from 0 and 1 uniformly in $p \in (\delta, 1-\delta)$ and in the configuration away from this edge. This property also extends to any finite family of edges. Hence,

$$\mathbb{P}^{(0)}_{p,q}\left(\bigcap_{i=1}^n \tau_{iLe}(A_{L,m}^c)\right) \leq \Big(1 - \mathbb{P}^{(0)}_{S(L,m),p,q}(A_{L,m})\Big)^n$$

Since $p > \widehat{p}_c(q, L)$,

$$\liminf_{m\to\infty} \mathbb{P}^{(0)}_{S(L,m),p,q}(A_{L,m}) > 0.$$

Hence for some $0 < a(L) < 1$,

$$\mathbb{P}^{(0)}_{p,q}\left(\bigcap_{i=1}^n \tau_{iLe}(A_L^c)\right) = \lim_{m\to\infty} \mathbb{P}^{(0)}_{p,q}\left(\bigcap_{i=1}^n \tau_{iLe}(A_{L,m}^c)\right) \leq a(L)^n \tag{4.14}$$

Then (4.13) and the above estimate imply that for $d \geq 3$, $\mathbb{P}_0\big\{|v_e| \geq n\big\}$ decays exponentially in n. To prove this statement in $d = 2$, we again recall the definition of the slab-critical probability $\widehat{p}_c(q)$ and note that $p > \widehat{p}_c(q)$. In this regime, we have exponential decay of truncated connectivity (see [G06, Theorem 5.108 and the following paragraph]). In other words, if $\mathcal{C}$ denotes an open cluster at the origin, then

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}^{(0)}_{p,q}\Big\{\big|\mathcal{C}\big| \geq n^2\ ;\big|\mathcal{C}\big| < \infty\Big\} < 0. \tag{4.15}$$

In this super-critical regime $p > p_c(q)$, we also have exponential decay of dual connectivity (see [BD12, Theorems 1 and 2]). In other words, in the dual random cluster model in $d = 2$, the probability for two points x and y to be connected by a path decays exponentially fast with respect to the distance between x and y.

To show that $\mathbb{P}_0\big\{|v_e| \geq n\big\}$ decays exponentially in n in $d = 2$, we now let B_n to be the box $\{1, \dots, n\} \times \{1, \dots, n\}$. Then on the event $\{|v_e| \geq n,\ ; 0 \in \mathcal{C}_\infty\}$, none of the boundary sites $\{je : j = 1, ..., n\}$ are in $\mathcal{C}_\infty$. Hence, either at least one of these sites is in a finite component of size larger than n or there exists a dual crossing of B_n in the direction of e. The probabilities of both these events are exponentially small in n by (4.15) and the exponential decay of dual connectivity. Hence $\mathbb{P}_0\big\{|v_e| \geq n\big\}$ decays exponentially in n for $d \geq 2$.

To conclude the proof of Lemma 4.3, we note that for any $\varepsilon > 0$,

$$\big\{\ell > n\big\} = \bigcup_{j=1}^{\lceil \varepsilon n\rceil} \bigg\{\mathrm{d_{ch}}(0, je) > n\,;\ 0, je \in \mathcal{C}_\infty\bigg\} \bigcup \bigg\{|v_e| \geq \lceil \varepsilon n\rceil\bigg\}. \tag{4.16}$$

Since $\mathbb{P}_0$-probabilities of the events in the first union are exponentially small by the uniform estimate (4.12) on the chemical distance $\mathrm{d_{ch}}(0, je)$, and $\mathbb{P}_0\big\{|v_e| \geq n\big\}$ also decays exponentially in n, we now invoke union of events bound and absorb the linear factor coming from the number of events in the exponential bound and end up with the proof of Lemma 4.3. □

We now turn to the proof of Theorem 4.1.

Proof of Theorem 4.1: Let us fix any $G \in \mathcal{G}_\infty$ and for any nearest neighbor occupied path $0 = x_0, \ldots, x_n = x$ in $\mathcal{C}_\infty$, let $V(\omega, x) = V_G(\omega, x) = \sum_{j=0}^{n-1} G(\tau_{x_j}, x_{j+1} - x_j)$ as defined in (4.6). Recall that we have to show

$$\lim_{n\to\infty} \max_{\substack{x\in\mathcal{C}_\infty \\ |x|\leq n}} \frac{|V(\omega, x)|}{n} = 0 \qquad \mathbb{P}_0 - \text{ a.s.} \tag{4.17}$$

Let us first make an observation based on the facts proved in Lemma 4.2 and Lemma 4.3. Indeed, with of $V(\omega, v_e)$ defined in (4.9), Lemma 4.3 and our uniform bound assumption (4.4) imply that $\mathbb{E}_0[|V(\omega, v_e)|] < \infty$. Furthermore, $\mathbb{E}_0[V(\omega, v_e)] = 0$ by our induced mean-zero assumption (4.10). If we now write $F(\omega) = V(\omega, v_e)$ and recall that $n_{k+1}(\omega, e) = n_k(\sigma_e\omega, e)$ from (4.3), then $V(\omega, v_e) = V(\omega, n_k(\omega, e)\, e) = \sum_{j=0}^{k-1} F \circ \sigma_e^j(\omega)$. Since the induced shift $\sigma_e : \Omega_0 \to \Omega_0$ is measure-preserving and ergodic, by Birkhoff's ergodic theorem,

$$\lim_{k\to\infty} \frac{1}{k} V(\omega, n_k(\omega, e)\, e) = 0 \qquad \mathbb{P}_0 - \text{ a.s.} \tag{4.18}$$

We now fix $\varepsilon > 0$ which is arbitrary. We claim that

$$\lim_{n\to\infty} \frac{1}{n^d} \sum_{\substack{x\in\mathcal{C}_\infty \\ |x|\leq n}} 1\!\!1_{\big\{|V(x,\omega)| > \varepsilon n\big\}} = 0 \qquad \mathbb{P}_0 - \text{ a.s.} \tag{4.19}$$

Actually (4.18) forms the core of the argument for the proof of (4.19). Indeed, given (4.19), the proof of the claim (4.18) for all the percolation models including long-range correlations introduced in Section 1.1 and Section 1.2, now closely follows the proof of [BB07, Theorem 5.4] deduced for i.i.d. Bernoulli percolation. In fact, the crucial estimate [BB07, (5.28)] can be proved using the *FKG inequality*, which is available also in the all the models appearing in Section 1.1 and Section 1.2. Recall that the FKG inequality asserts that for two increasing events A and B (i.e, events that are preserved by addition of open edges), $\mathbb{P}(A \cap B) \geq \mathbb{P}(A)\mathbb{P}(B)$. For the validity of this positive correlation inequality, we refer to [G06, Theorem 4.17] for the random cluster model, to [T09] for the random interlacement and its vacant set, and to [R15, Remark 1.4] for the level sets of Gaussian free fields. Hence, based on the assertion (4.19) we have just proved and using the FKG inequality, we can repeat the arguments of [BB07, Theorem 5.4] to prove the estimate [BB07, (5.28)] therein and thus deduce (4.19). Then, for any arbitrary $\varepsilon > 0$, (4.19) in particular implies that, for n large enough,

$$\sum_{\substack{x\in\mathcal{C}_\infty \\ |x|\leq n}} 1\!\!1_{\big\{|V(x,\omega)| > \varepsilon n\big\}} < \varepsilon n^d \qquad \mathbb{P}_0 - \text{ a.s.} \tag{4.20}$$

Let us make another observation based on Lemma 4.2. Recall that $\theta(p) > 0$ denotes the percolation density, i.e., $\theta(p)$ is the probability that 0 is in the infinite open cluster $\mathcal{C}_\infty$. Also, for any arbitrary $\varepsilon > 0$ as before, let us set

$$\delta = \frac{1}{2}\bigg(\frac{4\varepsilon}{\theta(p)}\bigg)^{\frac{1}{d}}. \tag{4.21}$$

Then by Lemma 4.2, for any $x, y \in \mathcal{C}_\infty$ with $|x| < n$, $|y| < n$ and $|x-y| < \delta n$,

$$\mathrm{d}_{\mathsf{ch}}(x,y) < \rho\delta n. \tag{4.22}$$

Finally, let us observe that, by Birkhoff's ergodic theorem,

$$\lim_{n\to\infty} \frac{1}{(2n+1)^d} \sum_{|x|\le n} 1\!\!1\big\{x \in \mathcal{C}_\infty\big\} = \theta(p) \qquad \mathbb{P}_0 - \text{ a.s.}$$

Hence for any fixed $\delta > 0$, for every n large enough and $\mathbb{P}_0$-almost surely, in a ball of radius δn in $\mathcal{C}_\infty \cap [-n,n]^d$ there are at least $\delta^d (2n)^d \frac{\theta}{2}$ points in $\mathcal{C}_\infty$. In particular, for our choice of δ as required in (4.21),

$$\#\big\{\text{points in a box of radius } \delta n \text{ in } [-n,n]^d \text{ in } \mathcal{C}_\infty\big\} > 2\varepsilon n^d. \tag{4.23}$$

Given (4.20) and (4.23), we now claim that, for large enough n and every $x \in [-n,n]^d$, there exists $y \in [-n,n]^d \cap \mathcal{C}_\infty$ so that $|y - x| < \delta n$ and

$$|V(\omega,y)| \le \varepsilon n \qquad \mathbb{P}_0 - \text{ a.s.}$$

Indeed, by (4.20) there are at most εn^d points $z \in [-n,n]^d$ such that $|V(\omega,z)| \ge \varepsilon n$ and by (4.23), there are at least $2\varepsilon n^d$ points in $B_{n\delta}(x) \cap \mathcal{C}_\infty$. Hence, we have at least one point $y \in [-n,n]^d \cap \mathcal{C}_\infty$ such that $|y-x| < \delta n$ and $|V(\omega,y)| \le \varepsilon n$, $\mathbb{P}_0$- almost surely.

Let us now prove (4.18). Recall the definition of V from (4.6). Then, by (4.22),

$$\begin{aligned}\big|V(\omega,x) - V(\omega,y)\big| &\le \mathrm{d}_{\mathsf{ch}}(x,y)\, \text{ess} \sup_{\omega-\mathbb{P}_0} G(\omega,x)\\ &\le \rho\delta n A,\end{aligned}$$

for some $A < \infty$, recall (4.4). Since $|V(\omega,y)| \le \varepsilon n$, then $\mathbb{P}_0$- almost surely,

$$\begin{aligned}|V(\omega,x)| &\le |V(\omega,y)| + \rho\delta n A\\ &\le \varepsilon n + \rho\delta n A.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary and $\delta \to 0$ as $\varepsilon \to 0$ according to (4.21), Theorem 4.1 is proved. □

We have an immediate corollary to Theorem 4.1.

Corollary 4.4. *Let $G \in \mathcal{G}_\infty$. For every $\varepsilon > 0$, there exists $c_\varepsilon = c_\varepsilon(\omega)$ so that, for every sequence of points $(x_k)_{k=0}^n$ on $\mathcal{C}_\infty$ with $x_0 = 0$ and $|x_{k+1} - x_k| = 1$,*

$$\Big|\sum_{k=0}^{n-1} G(\tau_{x_k}\omega, x_{k+1} - x_k)\Big| \le c_\varepsilon + n\varepsilon.$$

In particular,

$$\sum_{k=0}^{n-1} G(\tau_{x_k}\omega, x_{k+1} - x_k) \ge -c_\varepsilon - n\varepsilon. \tag{4.24}$$

4.3 Proof of the upper bound for Theorem 2.3. We now prove the upper bound in Theorem 2.3 using the sub-linear growth property of gradient functions established in Theorem 4.1 and Corollary 4.4.

Lemma 4.5 (The upper bound)**.** *For $\mathbb{P}_0$- almost every $\omega \in \Omega_0$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log E_0^{\pi,\omega}\bigg\{\exp\Big\{\sum_{k=0}^{n-1} f\big(\tau_{X_k}\omega, X_{k+1} - X_k\big)\Big\}\bigg\} \le \inf_{G\in\mathcal{G}_\infty} \Lambda(f,G),$$

where

$$\Lambda(f,G) = \operatorname{ess\,sup}_{\mathbb{P}_0} \bigg(\log \sum_e 1\!\!1_{\{\omega(e)=1\}} \pi_\omega(0,e) \exp\big\{f(\omega,e) + G(\omega,e)\big\}\bigg). \tag{4.25}$$

Proof. Fix $G \in \mathcal{G}_\infty$. By the definition of the Markov chain $P_0^{\pi,\omega}$ we have,

$$\begin{aligned}
&E_0^{\pi,\omega}\Big\{ \exp\Big\{ f(\tau_{X_k}\omega, X_{k+1}-X_k) + G(\tau_{X_{k+1}}\omega, X_{k+1}-X_k)\Big\}\Big| X_k\Big\} \\
&= \sum_{|e|=1} \pi_\omega\big(X_k, X_k+e\big) \mathrm{e}^{f(\tau_{X_k}\omega,e)+G(\tau_{X_k}\omega,e)} \\
&= \sum_{|e|=1} 1\!\!1_{\{(\tau_{X_k}\omega)(e)=1\}} \pi_\omega\big(X_k, X_k+e\big) \mathrm{e}^{f(\tau_{X_k}\omega,e)+G(\tau_{X_k}\omega,e)} \\
&\le \mathrm{e}^{\Lambda(f,G)},
\end{aligned}$$

where the uniform upper bound follows from (4.25).

Invoking the Markov property and successive conditioning, we have

$$E_0^{\pi,\omega}\Big\{ \exp\Big\{ \sum_{k=0}^{n-1} \Big(f(\tau_{X_k}\omega, X_{k+1}-X_k) + G(\tau_{X_k}\omega, X_{k+1}-X_k)\Big)\Big\}\Big\} \le \mathrm{e}^{n\Lambda(f,G)}. \tag{4.26}$$

We now recall Corollary 4.4 and plug in the lower bound (4.24) in (4.26). Then if we divide both sides by n, take logarithm and pass to $\limsup_{n\to\infty}$, we obtain the upper bound

$$\limsup_{n\to\infty} \frac{1}{n} \log E_0^{\pi,\omega}\Big\{ \exp\big\{ \sum_{k=0}^{n-1} f\big(\tau_{X_k}\omega, X_{k+1}-X_k\big)\big\}\Big\} \le \Lambda(f,G) + \varepsilon.$$

Passing to $\varepsilon \to 0$ and subsequently taking $\inf_{G\in\mathcal{G}_\infty}$ we finish the proof of the lemma. ☐

5. Equivalence of bounds: Min-max Theorems based on entropic coercivity

In this section we turn to the proof of the crucial fact that the lower bound obtained from Corollary 3.5 and the upper bound from Lemma 4.5 indeed match. The following theorem holds the key argument of our analysis and will also finish the proof of Theorem 2.3. Recall the lower bound variational formula $\overline{H}(f)$ from (3.14), and the upper bound variational formula $\Lambda(f,G)$ from (4.25).

Theorem 5.1 (Equivalence of bounds)**.** *For any continuous and bounded function f on $\Omega_0 \times \mathbb{B}_d$,*

$$\begin{aligned}
\overline{H}(f) &= \inf_{G\in\mathcal{G}_\infty} \operatorname{ess\,sup}_{\mathbb{P}_0} \Big(\log \sum_{e\in\mathbb{B}_d} 1\!\!1_{\omega(e)=1}\, \pi_\omega(0,e)\, \exp\big\{f(\omega,e)+G(\omega,e)\big\}\Big) \\
&= \inf_{G\in\mathcal{G}_\infty} \Lambda(f,G)
\end{aligned}$$

We will prove Theorem 5.1 in several steps. The first step is to invoke a min-max argument to exchange the order of $\sup_{\widetilde{\pi}}$ and $\inf_g$ in (5.2), and subsequently solve the maximization problem in $\widetilde{\pi}$. The resulting assertion is

Lemma 5.2 (Entropic coercivity in $\widetilde{\pi}$)**.** *For any continuous and bounded function f on $\Omega_0 \times \mathbb{B}_d$,*

$$\overline{H}(f) = \sup_\phi \inf_g \int \mathrm{d}\mathbb{P}_0(\omega)\, \phi(\omega) L(g,\omega)$$

where,

$$L(g,\omega) = L_f(g,\omega) = \log\Big(\sum_{e\in\mathbb{B}_d} \pi_\omega(0,e) \exp\big\{f(\omega,e)+g(\omega)-g(\tau_e\omega)\big\}\Big). \tag{5.1}$$

Proof. Let us denote

$$\overline{H}(f) = \sup_{\phi} \sup_{\widetilde{\pi}\in\widetilde{\Pi}} \inf_{g} \bigg[\int \mathrm{d}\mathbb{P}_0(\omega)\, \phi(\omega) \bigg\{ \sum_{e\in\mathbb{B}_d} \widetilde{\pi}(\omega,e)\big(F(\omega,e) - \log\widetilde{\pi}(\omega,e)\big) \bigg\} \bigg], \tag{5.2}$$

where

$$F(\omega,e) = F(\pi,f,g,\omega,e) = f(\omega,e) + \log\pi_\omega(0,e) + \big(g(\omega) - g(\tau_e\omega)\big). \tag{5.3}$$

For any $\widetilde{\pi}$ and g, let us write the functional

$$\begin{aligned} \mathfrak{F}(\widetilde{\pi},g) &= \int \mathrm{d}\mathbb{P}_0(\omega)\, \phi(\omega) \bigg\{ \sum_{e\in\mathbb{B}_d} \widetilde{\pi}(\omega,e)\big[f(\omega,e) + \log\pi_\omega(0,e) + \big(g(\omega)-g(\tau_e\omega)\big) - \log\widetilde{\pi}(\omega,e)\big] \bigg\} \\ &= \int \mathrm{d}\mathbb{P}_0(\omega)\, \phi(\omega) \bigg\{ \sum_{e\in\mathbb{B}_d} \widetilde{\pi}(\omega,e)\big[F(\omega,e) - \log\widetilde{\pi}(\omega,e)\big] \bigg\} \end{aligned} \tag{5.4}$$

with $F(\omega,e)$ defined in (5.3) (recall (5.2)). First we would like to show that, for any fixed ϕ,

$$\sup_{\widetilde{\pi}\in\widetilde{\Pi}} \inf_{g} \mathfrak{F}(\widetilde{\pi},g) = \inf_{g} \sup_{\widetilde{\pi}\in\widetilde{\Pi}} \mathfrak{F}(\widetilde{\pi},g). \tag{5.5}$$

This requires the following coercivity argument. Note that, for each fixed ϕ, and corresponding to any $\widetilde{\pi}\in\widetilde{\Pi}$, we have the entropy functional

$$\mathrm{Ent}(\mu_{\widetilde{\pi}}) = \int \sum_{e} \widetilde{\pi}(\omega,e)\log\widetilde{\pi}(\omega,e)\phi(\omega)\mathrm{d}\mathbb{P}_0(\omega)$$

corresponding to the probability measure $\mathrm{d}\mu_{\widetilde{\pi}}(\omega,e) = \widetilde{\pi}(\omega,e)\big(\phi(\omega)\mathrm{d}\mathbb{P}_0(\omega)\big) \in \mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$. Then for any fixed ϕ, the map $\widetilde{\pi}\mapsto\mathrm{Ent}(\mu_{\widetilde{\pi}})$ is convex, lower semi-continuous and has weakly compact level sets (i.e., for any $a\in\mathbb{R}$, the set $\{\widetilde{\pi}\in\widetilde{\Pi}\colon \mathrm{Ent}(\mu_{\widetilde{\pi}})\le a\}$ is weakly compact). Furthermore, for any probability density ϕ, any continuous and bounded function f on $\Omega_0\times\mathbb{B}_d$ and bounded measurable function g, and for every $\widetilde{\pi}\in\widetilde{\Pi}$,

$$\begin{aligned} \int \phi(\omega) \sum_{e\in\mathbb{B}_d} \widetilde{\pi}(\omega,e)F(\omega,e)\mathrm{d}\mathbb{P}_0 &= \int \phi(\omega) \sum_{e\in\mathbb{B}_d} \widetilde{\pi}(\omega,e)\big[f(\omega,e) + \log\pi_\omega(0,e) + \big(g(\omega)-g(\tau_e\omega)\big]\mathrm{d}\mathbb{P}_0 \\ &\le \big(\|f\|_\infty + 2\|g\|_\infty\big) \int \phi(\omega) \sum_{e\in\mathbb{B}_d} \widetilde{\pi}(\omega,e)\mathrm{d}\mathbb{P}_0 \\ &= \|f\|_\infty + 2\|g\|_\infty := C < \infty. \end{aligned}$$

We conclude that for any g, the map $\widetilde{\pi}\mapsto\mathfrak{F}(\widetilde{\pi},g)$ is concave, weakly upper-semicontinuous and has weakly compact "upper level sets"$\{\widetilde{\pi}\colon \mathfrak{F}(\widetilde{\pi},g)\ge a\}$ for any $a\in\mathbb{R}$. Furthermore, for any $\widetilde{\pi}\in\widetilde{\Pi}$, the map $g\mapsto\mathfrak{F}(\widetilde{\pi},g)$ is linear and continuous in g. Then, in view of *Von-Neumann's min-max theorem* (p. 319, [AE84]), the equality (5.5) holds. Hence,

$$\overline{H}(f) = \sup_{\phi} \inf_{g} \sup_{\widetilde{\pi}} \bigg[\int \mathrm{d}\mathbb{P}_0(\omega)\, \phi(\omega) \bigg\{ \sum_{e\in\mathbb{B}_d} \widetilde{\pi}(\omega,e)\big[F(\omega,e) - \log\widetilde{\pi}(\omega,e)\big] \bigg\} \bigg] \tag{5.6}$$

Since the integrand above depends only locally in $\widetilde{\pi}$, we can bring the $\sup_{\widetilde{\pi}\in\widetilde{\Pi}}$ inside the integral, and solve the variational problem

$$\sup_{\widetilde{\pi}} \sum_{e\in\mathbb{B}_d} \widetilde{\pi}(\omega,e)\big[F(\omega,e) - \log\widetilde{\pi}(\omega,e)\big]$$

subject to the Lagrange multiplier constraint $\sum_e \widetilde{\pi}(\cdot,e) = 1$. The maximizer is

$$\widetilde{\pi}(\cdot,e) = \frac{\exp[F(\omega,e)]}{\sum_{e\in\mathbb{B}_d}\exp[F(\omega,e)]},$$

and if we plug in this value in (5.6) and recall the definition of $F(\omega, e)$ from (5.3), then (5.6) leads us to

$$\begin{aligned}\overline{H}(f) &= \sup_{\phi} \inf_{g} \bigg[\int \mathrm{d}\mathbb{P}_0(\omega)\, \phi(\omega) \log \bigg(\sum_{e \in \mathbb{B}_d} \pi_\omega(0,e) \exp\big\{ f(\omega,e) + g(\omega) - g(\tau_e \omega) \big\} \bigg) \bigg] \\ &= \sup_{\phi} \inf_{g} \int \mathrm{d}\mathbb{P}_0(\omega)\, \phi(\omega) L(g,\omega),\end{aligned} \tag{5.7}$$

which concludes the proof of Lemma 5.2. □

Now we would like to exchange $\sup_\phi$ and $\inf_g$ in (5.7). For this, we need to invoke a *compactification* argument based on an *entropy penalization method*. This is the the content of the next lemma.

Lemma 5.3 (Entropy penalization and coercivity in ϕ)**.** *For any continuous and bounded function f on $\Omega_0 \times \mathbb{B}_d$,*

$$\overline{H}(f) \geq \liminf_{\varepsilon \to 0} \inf_{g} \varepsilon \log \mathbb{E}_0 \Big[\mathrm{e}^{\varepsilon^{-1} L(g,\cdot)} \Big]$$

where $L(g,\cdot)$ is the functional defined in (5.1)*.*

Proof. We start from (5.7). For any probability density $\phi \in L^1_+(\mathbb{P}_0)$, note that its entropy functional

$$\mathrm{Ent}(\phi) = \int \phi(\omega)\, \log \phi(\omega)\, \mathrm{d}\mathbb{P}_0(\omega).$$

is always non-negative by Jensen's inequality. Hence, for any fixed $\varepsilon > 0$, we have a lower bound

$$\overline{H}(f) \geq \sup_{\phi} \inf_{g} \bigg[\int \mathrm{d}\mathbb{P}_0(\omega)\, \phi(\omega) \Big(L(g,\omega) - \varepsilon \log \phi(\omega) \Big) \bigg]. \tag{5.8}$$

Again, $\phi \mapsto \mathrm{Ent}(\phi)$ is convex and weakly lower semicontinuous in $L^1_+(\mathbb{P}_0)$, with its level sets $\{\phi\colon \int \phi \log \phi \mathrm{d}\mathbb{P}_0 \leq a\}$ being weakly compact in $L^1_+(\mathbb{P}_0)$ for all $a \in \mathbb{R}$. Also, by (5.1), for any bounded f on $\Omega_0 \times \mathbb{B}_d$ and bounded g, and for every ϕ,

$$\int \phi(\omega) L(g,\omega) \mathrm{d}\mathbb{P}_0 = \int \phi(\omega) \log \bigg(\sum_{e \in \mathbb{B}_d} \pi_\omega(0,e) \exp\big\{ f(\omega,e) + g(\omega) - g(\tau_e\omega) \big\} \bigg) \leq \|f\|_\infty + 2\|g\|_\infty = C < \infty.$$

Then, if we write

$$\mathcal{A}_\varepsilon(g,\phi) = \int \mathrm{d}\mathbb{P}_0(\omega)\, \big[\phi(\omega) L(g,\omega) - \varepsilon \phi(\omega) \log \phi(\omega) \big], \tag{5.9}$$

then, for every $\varepsilon > 0$, like in Lemma 5.2, the map $g \mapsto \mathcal{A}_\varepsilon(g,\phi)$ is convex and continuous and the map $\phi \mapsto \mathfrak{H}_\varepsilon(g,\phi)$ is concave and upper semicontinuous with compact "upper level sets"(i.e. the set $\{\phi\colon \mathcal{A}_\varepsilon(g,\phi) \geq a\}$ is weakly compact for all $a \in \mathbb{R}$). Applying Von-Neumann's min-max theorem once more, we can swap the order of $\sup_\phi$ and $\inf_g$ in (5.8). Hence,

$$\begin{aligned}\overline{H}(f) \geq \inf_{g} \sup_{\phi} \mathcal{A}_\varepsilon(g,\phi) &= \inf_{g} \varepsilon \log \mathbb{E}_0 \Big[\mathrm{e}^{\varepsilon^{-1} L(g,\cdot)} \Big] \\ &\geq \liminf_{\varepsilon \to 0} \inf_{g} \varepsilon \log \mathbb{E}_0 \Big[\mathrm{e}^{\varepsilon^{-1} L(g,\cdot)} \Big].\end{aligned} \tag{5.10}$$

We remark that the second identity above follows from a standard perturbation argument in ϕ and the definition of $\mathcal{A}_\varepsilon$ set in (5.9). Indeed, for any admissible class of test functions ψ, we need to solve for ϕ by setting

$$\frac{\mathrm{d}}{\mathrm{d}\eta}\bigg|_{\eta=0} \Big[\mathcal{A}_\varepsilon(g, \phi + \eta\psi) \Big] = 0$$

for any fixed $\varepsilon > 0$ and g, and subject to the condition $\int \phi \mathrm{d}\mathbb{P}_0 = 1$. The solution is given by

$$\phi(\cdot) = \frac{\exp\{\varepsilon^{-1} L(g, \cdot)\}}{\mathbb{E}_0[\exp\{\varepsilon^{-1} L(g, \cdot)\}]}.$$

If we substitute this value of ϕ in (5.9), then we are led to the identity (5.10). This concludes the proof of Lemma 5.3. ☐

We need the following important lemma, whose proof is deferred to until the end of the proof of Theorem 5.1. Recall that $\mathbb{U}_d = \{\pm u_i\}_{i=1}^d$ the nearest neighbors of the origin 0.

Lemma 5.4. *There exists a sequence $\varepsilon_n \to 0$ and a sequence g_n of bounded measurable functions such that,*

$$\overline{H}(f) \geq \varepsilon_n \log \mathbb{E}_0 \Big[\mathrm{e}^{\varepsilon_n^{-1} L(g_n, \cdot)} \Big], \tag{5.11}$$

and for any $u \in \mathbb{U}_d$,

$$G_n(\omega, u) = \mathbb{1}\{0 \in \mathcal{C}_\infty\} \, \mathbb{1}\{\omega(u) = 1\} \, \big(g_n(\omega) - g_n(\tau_u \omega) \big) \tag{5.12}$$

converges weakly along some subsequence to some $G(\cdot, u)$. Furthermore, $G \in \mathcal{G}_\infty$.

We first assume the above lemma and prove Theorem 5.1. For this purpose, we need another lemma.

Lemma 5.5. *For any $\lambda > 0$, any probability measure μ and for any random variable X with finite exponential moment, if we set*

$$\psi(\lambda) = \log \mathbb{E}^{(\mu)} \big[\mathrm{e}^{\lambda X} \big],$$

then the map $\lambda \mapsto \frac{\psi(\lambda)}{\lambda}$ is increasing in $[0, \infty)$.

Proof. Indeed, $\psi(\lambda)$ is convex and twice differentiable in λ. In particular, $\psi''(\lambda) > 0$, $\psi(0) = 0$ and

$$\left(\frac{\psi(\lambda)}{\lambda} \right)' = \frac{\psi'(\lambda)}{\lambda} - \frac{\psi'(\lambda)}{\lambda^2} = \frac{\lambda \psi'(\lambda) - \psi(\lambda)}{\lambda^2}.$$

Since $\lambda \psi'(\lambda) - \psi(\lambda)$ is 0 at $\lambda = 0$ and $(\lambda \psi'(\lambda) - \psi(\lambda))' = \lambda \psi'' > 0$, we conclude that $\lambda \mapsto \frac{\psi(\lambda)}{\lambda}$ is increasing in $\lambda > 0$. ☐

We now continue with the proof of Theorem 5.1 assuming Lemma 5.4.

Proof of Theorem 5.1: Note that Lemma 5.5 and (5.11) imply that for $\lambda > 0$ and large enough n,

$$\overline{H}(f) \geq \frac{1}{\lambda} \log \mathbb{E}_0 \big[\mathrm{e}^{\lambda L(g_n, \cdot)} \big].$$

Plugging in the expression for $L(g_n, \cdot)$ from (5.1) and recalling the definition of G_n from (5.12), we then get,

$$\overline{H}(f) \geq \frac{1}{\lambda} \log \mathbb{E}_0 \bigg[\exp \bigg\{ \lambda \log \bigg(\sum_{e \in \mathbb{B}_d} \pi_\omega(0, e) \exp \big\{ f(\omega, e) + G_n(\omega, e) \big\} \bigg) \bigg\} \bigg]$$

If we now let $n \to \infty$, the first part of Lemma 5.4 implies that $G_n(\omega, e)$ converges weakly to some $G(\omega, e)$. Hence,

$$\overline{H}(f) \geq \frac{1}{\lambda} \log \mathbb{E}_0 \bigg[\exp \bigg\{ \lambda \log \bigg(\sum_{e \in \mathbb{B}_d} \pi_\omega(0, e) \exp \big\{ f(\omega, e) + G(\omega, e) \big\} \bigg) \bigg\} \bigg]. \tag{5.13}$$

If we now let $\lambda \to \infty$, we deduce that

$$\begin{aligned}\overline{H}(f) &\geq \operatorname*{ess\,sup}_{\mathbb{P}_0} \log\bigg(\sum_{e\in\mathbb{B}_d} \pi_\omega(0,e)\exp\big\{f(\omega,e)+G(\omega,e)\big\}\bigg) \\ &\geq \inf_{G\in\mathcal{G}_\infty} \operatorname*{ess\,sup}_{\mathbb{P}_0} \log\bigg(\sum_{e\in\mathbb{B}_d} \pi_\omega(0,e)\exp\big\{f(\omega,e)+G(\omega,e)\big\}\bigg),\end{aligned} \tag{5.14}$$

and in the last lower bound we invoked the second part of Lemma 5.4 which asserts that $G \in \mathcal{G}_\infty$. This proves Theorem 5.1, assuming Lemma 5.4. □

We now the owe the reader the proof of Lemma 5.4.

Proof of Lemma 5.4: We will prove Lemma 5.4 in two main steps. In the first step we will show that the sequence of formal gradients G_n defined in (5.12) is uniformly integrable and converges along a subsequence to some G. In the next step we will show that the limit G belongs to the class $\mathcal{G}_\infty$ introduced in Section 4.1.

Step 1: Proving uniform integrability of G_n. First we want to prove that G_n defined in (5.12) is uniformly integrable. Note that by (5.10), there exists a $\varepsilon_n \to 0$ and a sequence $(g_n)_n$ of bounded measurable functions so that

$$\varepsilon_n \log \mathbb{E}_0\bigg[\exp\bigg\{\varepsilon_n^{-1}\log\bigg(\sum_{e\in\mathbb{B}_d}\pi_\omega(0,e)\exp\big\{f(\omega,e)+g_n(\omega)-g_n(\tau_e\omega)\big\}\bigg)\bigg\}\bigg] \leq \overline{H}(f).$$

Since f is bounded, $f(\omega,e) \geq -\|f\|_\infty$ and by Lemma 5.5, in particular we have

$$\mathbb{E}_0\bigg[\sum_{e\in\mathbb{B}_d}\pi_\omega(0,e)\exp\big\{g_n(\omega)-g_n(\tau_e\omega)\big\}\bigg] \leq \exp\{\overline{H}(f)+\|f\|_\infty\} \tag{5.15}$$

Recall that $\mathbb{U}_d = \{\pm u_i\}_{i=1}^d$ the nearest neighbors of the origin 0. For any $u = \pm u_i$, let $\Omega_{0,u}$ denote the set of configurations ω such that both 0 and u are in the infinite cluster $\mathcal{C}_\infty(\omega)$ and the edge $0 \leftrightarrow u$ is present (i.e., $\omega(u) = 1$). Then $\mathbb{P}(\Omega_{0,u}) > 0$ and we set $\mathbb{P}_{0,u}(\cdot) = \mathbb{P}(\cdot|\Omega_{0,u})$.

Now for any $u = \pm u_i$, if the edge $0 \leftrightarrow u$ is present, then $\pi_\omega(0,u) \geq 1/2d > 0$ and for some constant $C > 0$, (5.15) implies

$$\mathbb{E}_{0,u}\big[\exp\big\{g_n(\omega)-g_n(\tau_u\omega)\big\}\big] \leq C. \tag{5.16}$$

Now again by (5.15),

$$\begin{aligned}\mathbb{E}_{0,u}\bigg[\sum_{e\in\mathbb{B}_d}\pi_{\tau_u\omega}(0,e)\exp\big\{g_n(\tau_u\omega)-g_n(\tau_e\tau_u\omega)\big\}\bigg] &= \mathbb{E}_0\bigg[\sum_{e\in\mathbb{B}_d}\pi_\omega(0,e)\exp\big\{g_n(\omega)-g_n(\tau_e\omega)\big\}\bigg] \\ &\leq \exp\{\overline{H}(f)+\|f\|_\infty\}\end{aligned}$$

Now if the edge $0 \leftrightarrow u$ is present in the configuration ω (i.e., $\omega(u) = 1$), the edge $-u \leftrightarrow 0$ is present in the configuration $\tau_u\omega$ (i.e., $\pi_{\tau_u\omega}(0,-u) \geq 1/2d > 0$) and hence, again

$$\mathbb{E}_{0,u}\big[\exp\big\{g_n(\tau_u\omega)-g_n(\tau_{-u}\tau_u\omega)\big\}\big] = \mathbb{E}_{0,u}\big[\exp\big\{g_n(\tau_u\omega)-g_n(\omega)\big\}\big] \leq C. \tag{5.17}$$

It follows from (5.16) and (5.17) that the sequence G_n defined in (5.12) is uniformly integrable under $\mathbb{P}_0$. Hence it is also uniformly tight and converges weakly along a subsequence to some G.

Step 2: Proving that $G \in \mathcal{G}_\infty$. To conclude that $G \in \mathcal{G}_\infty$, note that clearly G satisfies the closed loop property (4.5) on the infinite cluster $\mathcal{C}_\infty$ as G_n is a gradient field on the infinite cluster $\mathcal{C}_\infty$. Furthermore, the fact that G is bounded in the essential supremum norm in $\mathbb{P}_0$ follows again from the first inequality in the display (5.14)[1]

[1] Note that the display (5.14) followed only from the first part of Lemma 5.4 (i.e., the fact that that G_n converges weakly along a subsequence to some G), which we have just proved in Step 1. In particular, (5.14) does not use the second part of Lemma 5.4 which asserts that $G \in \mathcal{G}_\infty$, which we are proving currently in Step 2.

It remains to check the induced zero mean property (4.10) of G. The following lemma will finish the proof of Lemma 5.4. Hence, the proof of Theorem 5.1 will also be concluded.

Lemma 5.6 (Induced mean zero property of the limit G)**.** *The limiting gradient G appearing in Lemma 5.4 satisfies the induced mean zero property defined in* (4.10)*. Hence, $G \in \mathcal{G}_\infty$.*

Proof. Let us fix $e \in \mathbb{B}_d$ and recall that ℓ denotes the graph distance from 0 to $v_e = k(\omega, e)e$, and fix $(x_0 = 0, x_1, \ldots, x_\ell)$ a shortest open path to $v_e = k(\omega, e)e$. We also recall from Section 4 that the induced shift $\sigma_e \colon \Omega_0 \to \Omega_0$ defined by $\sigma_e(\omega) = \tau_{k(\omega,e)e}(\omega)$ is $\mathbb{P}_0$-measure preserving. Hence, for any bounded measurable g_n

$$\mathbb{E}_0\big[g_n(\tau_{k(\omega,e)e}\omega) - g_n(\omega)\big] = \mathbb{E}_0\big[g_n \circ \sigma_e - g_n\big] = 0. \tag{5.18}$$

Let us write,

$$F_M = \mathbb{E}_0\bigg[V(\omega,\, k(\omega,e)e)\, 1\!\!1_{\ell<M}\bigg] = \mathbb{E}_0\bigg[\sum_{j=0}^{\ell-1} G\big(\tau_{x_i}\omega, x_{i+1} - x_i\big)\ 1\!\!1_{\ell<M}\bigg].$$

We claim that,

$$F_M \to 0 \quad \text{as} \ \ M \to \infty. \tag{5.19}$$

Note that, by (5.12),

$$F_M = \mathbb{E}_0\bigg[1\!\!1_{\ell\le M} \sum_{j=0}^{\ell-1} \lim_{n\to\infty} G_n(\tau_{x_j}\omega, x_{j+1} - x_j)\bigg] = \mathbb{E}_0\bigg[1\!\!1_{\ell<M}\ \lim_{n\to\infty} \big(g_n(\tau_{k(\omega,e)e}\omega) - g_n(\omega)\big)\bigg],$$

and furthermore, by (5.18),

$$\big|F_M\big| \le \liminf_{n\to\infty} \bigg|\mathbb{E}_0\bigg[1\!\!1_{\ell\ge M}\big(g_n(\tau_{k(\omega,e)e}\omega) - g_n(\omega)\big)\bigg]\bigg|.$$

But

$$\begin{aligned}
\bigg|\mathbb{E}_0\bigg[1\!\!1_{\ell\ge M}\Big(g_n(\tau_{k(\omega,e)e}\omega) - g_n(\omega)\Big)\bigg]\bigg| &= \bigg|\mathbb{E}_0\bigg[\sum_{j=M}^{\infty} 1\!\!1_{\ell=j}\Big(\sum_{i=0}^{j-1}\big(g_n(\tau_{x_i}\omega) - g_n(\tau_{x_{i+1}}\omega)\big)\Big)\bigg]\bigg| \\
&= \bigg|\mathbb{E}_0\bigg[\sum_{j=M}^{\infty} 1\!\!1_{\ell=j}\Big(\sum_{i=0}^{j-1} G_n(\tau_{x_i}\omega, x_{i+1} - x_i)\Big)\bigg]\bigg| \\
&\le K\ \mathbb{E}_0\big[\sum_{j=M}^{\infty} j 1\!\!1_{\ell=j}\big],
\end{aligned}$$

and the last inequality follows from (5.16) and (5.17) for some constant $K > 0$. But by Lemma 4.3 the last term in the above display goes to 0 as $M \to \infty$. This proves the claim (5.19).

Recall the definition of the corrector $V(\omega, k(\omega, e)e) = \sum_{i=0}^{\ell-1} G(\tau_{x_i}\omega, x_{i+1} - x_i)$ corresponding to the limit G of G_n. To prove the induced mean zero property (4.10) for V, we have to show that $\mathbb{E}_0\big[V(\omega, k(\omega, e)e)\big] = 0$. For this, we note that $V(\omega, k(\omega, e)e)$ is the almost sure pointwise limit of $V(\omega, k(\omega, e)e)1\!\!1_{\ell\le M}$ as $M \to \infty$. Furthermore, $|V(\omega, k(\omega, e)e)1\!\!1_{\ell\le M}| \le \|G\|_\infty \ell$ for all M and $\mathbb{E}_0(\ell) < \infty$ by Lemma 4.3, so by the dominated convergence theorem

$$\mathbb{E}_0\big[V(\omega, k(\omega, e)e)\big] = \lim_{n\to\infty} \mathbb{E}_0\big[V(\omega, k(\omega, e)e)\, 1\!\!1_{\ell\le M}\big] = \lim_{M\to\infty} F_M = 0.$$

We conclude that G satisfies the induced mean zero property defined in (4.10). Hence, $G \in \mathcal{G}_\infty$ and the proofs of Lemma 5.6 and that of Lemma 5.4 are finished. This also concludes the proof of Theorem 5.1. $\square$

6. Proofs of Theorem 2.3, Theorem 2.1, Corollary 2.4 and Lemma 2.2

Proof of Theorem 2.3: The proof of Theorem 2.3 is readily finished by combining the lower bound from Corollary 3.5, the upper bound from Lemma 4.5 and the equivalence of bounds from Theorem 5.1. □

Proof of Theorem 2.1: By Theorem 2.3,

$$\lim_{n\to\infty}\frac{1}{n}\log E_0^{\pi,\omega}\big\{\exp\big\{n\langle f,\mathfrak{L}_n\rangle\big\}\big\}=\sup_{\mu\in\mathcal{M}_1^\star}\big\{\langle f,\mu\rangle-\mathfrak{I}(\mu)\big\}=\sup_{\mu\in\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)}\big\{\langle f,\mu\rangle-\mathfrak{I}(\mu)\big\}=\mathfrak{I}^\star(\mu).$$

Since Ω_0 is a closed subset of $\Omega=\{0,1\}^{\mathbb{B}_d}$ and hence, is compact, $\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$ is compact in the weak topology. The upper bound (2.5) for all closed sets now follows from Theorem 4.5.3 [DZ98]. The lower bound (2.6) has been proved by Lemma 3.4. □

Proof of Corollary 2.4: The claim follows by contraction principle once we show that $\inf_{\xi(\mu)=x}\mathfrak{I}(\mu) = \inf_{\xi(\mu)=x}\mathfrak{I}^{\star\star}(\mu)$. This is easy to check using convexity of $\mathfrak{I}$ and $\mathfrak{I}^{\star\star}$. □

Proof of Lemma 2.2: The zero speed regime of SRWPC under a drift.

For any $\beta>1$, we define

$$\pi^{(\beta)}(\omega,e)=\frac{V(e)1\!\!1_{\{\omega(e)=1\}}}{\sum_{e'\in\mathbb{B}_d}V(e')1\!\!1_{\{\omega(e')=1\}}}\in\widetilde{\Pi},$$

where

$$V(e)=\begin{cases}\beta>1 & \text{if } e=e_1,\\ 1 & \text{else.}\end{cases}$$

Let $X_n^{(\beta)}$ be the Markov chain with transition probabilities $\pi^{(\beta)}$. By [BGP03] and [S03], there exists $\beta_u=\beta_u(p,d)>0$ so that for $\beta>\beta_u$, the limiting speed

$$\lim_{n\to\infty}\frac{X_n^{(\beta)}}{n},$$

which exists and is an almost sure constant is zero. For the Bernoulli (bond and site) percolation cases the last statement follows from [BGP03] and [S03, Theorem 4.1]. Since the finite energy property (recall the proof of Lemma 4.3 for the random cluster model) holds for the random-cluster model [G06] and level sets of Gaussian free field [RS13, Remark 1.6], the proof of [S03, Theorem 4.1] is applicable, while in the case of random interlacements (for which the finite energy property fails), the statement regarding the zero speed of the random walk $X_n^{(\beta)}$ follows from [FP16].

Then, by Kesten's lemma (see [K75]), there exists no $\phi\in L^1(\mathbb{P}_0)$ so that $(\pi^{(\beta)},\phi)\in\mathcal{E}$. We split the proof into two cases.

Suppose there exists a neighborhood $\mathfrak{u}$ of $\pi^{(\beta)}$ so that every $\tilde{\pi}^{(\beta)}\in\overline{\mathfrak{u}}$ fails to have an invariant density. Then, for any $\tilde{\pi}^{(\beta)}\in\overline{\mathfrak{u}}$ and any probability density $\phi\in L^1(\mathbb{P}_0)$, let μ_β be the corresponding element in $\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$ (i.e., $\mathrm{d}\mu_\beta(\omega,e)=\pi^{(\beta)}(\omega,e)\phi(\omega)\mathrm{d}\mathbb{P}_0(\omega)$). Since

$$(\tilde{\pi}^{(\beta)},\phi)\notin\mathcal{E},$$

by Lemma 3.1, $\mu_\beta\notin\mathcal{M}_1^\star$. Then, $\mathfrak{I}(\mu_\beta)=\infty$ by (2.4). If $\mathfrak{I}$ were lower semicontinuous on $\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$, then $\mathfrak{I}=\mathfrak{I}^{\star\star}$ and by Theorem 2.1,

$$P_0^{\pi,\omega}\big\{\mathfrak{L}_n\in\mathfrak{n}\big\}\tag{6.1}$$

would decay super-exponentially for $\mathbb{P}_0$- almost every $\omega\in\Omega_0$, with $\mathfrak{n}$ being some neighborhood of μ_β. However, since for every ω, the relative entropy of $\pi^{(\beta)}(\omega,\cdot)$ w.r.t. $\pi_\omega(0,\cdot)$ is bounded below and above, the probability in (6.1) decays exponentially and we have a contradiction.

Assume that there exists no such neighborhood $\mathfrak{u}$ of $\pi^{(\beta)}$. Let $\tilde{\pi}_n\to\pi^{(\beta)}$ such that for all $n\in\mathbb{N}$, $\tilde{\pi}_n$ has an invariant density ϕ_n and $(\tilde{\pi}_n,\phi_n)\in\mathcal{E}$. If $(\mu_n)_n$ is the sequence corresponding to $(\tilde{\pi}_n,\phi_n)$, since $\mathcal{M}_1(\Omega_0\times\mathbb{B}_d)$ is compact, $\mu_n \Rightarrow \mu_\beta$ weakly along a subsequence. However, by our choice of $\beta > \beta_u$, $(\pi^{(\beta)}, \phi) \notin \mathcal{E}$ for any density ϕ and hence $\mu_\beta \notin \mathcal{M}_1^\star$ and $\mathfrak{I}(\mu_\beta) = \infty$. But,

$$\lim_{n\to\infty} \mathfrak{I}(\mu_n) = \int \mathrm{d}\mathbb{P}_0 \phi(\omega) \sum_{e\in\mathbb{B}_d} \pi^{(\beta)}(\omega, e) \log \frac{\pi^{(\beta)}(\omega, e)}{\pi_\omega(0, e)},$$

which is clearly finite. This proves that $\mathfrak{I}$ is not lower semicontinuous. □

References

[AE84] AUBIN and EKELAND Applied nonlinear analysis, Pure Appl. Math. (N. Y.), A WileyÐInterscience Publication, John Wiley and Sons, Inc., New York, 1984.

[AP96] P. ANTAL and A. PISZTORA. On the chemical distance for supercritical Bernoulli percolation, *Ann. Probab.* **24**, no. 2, 1036-1048, (1996)

[AS12] S. ARMSTRONG and P. SOUGANIDIS Stochastic homogenization of HamiltonÐJacobi and degenerate Bellman equations in unbounded environments, *Journal de Mathematiques Pures et Appliquees*, **97**, 5, 460-504, 2012

[BD12] V. BEFFARA and H. DUMINIL-COPIN. The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$, *Probab. Theory Related Fields*, **153** (2012), 511-542.

[BB07] N. BERGER and M. BISKUP. Quenched invariance principle for random walk on percolation clusters, *Probab. Theory Rel. Fields*, **137**, Issue 1-2, 83-120, (2007)

[BGP03] N. BERGER and N. GANTERT and Y. PERES. The speed of biased random walk on percolation clusters, *Probab. Theory Rel. Fields*, **126**, no. 2, 221–242, (2003)

[B05] T. BODINEAU. Slab percolation for the Ising model, Probab. Theory and Relat. Fields 132 (2005) 83-118.

[BK89] R. M. BURTON and M. KEANE, Density and uniqueness in percolation, *Comm. Math. Phys.*, **121**, no. 3, 501-505, (1989)

[CGZ00] F. COMETS, N. Gantert and O. Zeitouni, Quenched, annealed and functional large deviations for one dimensional random walks in random environments, Prob. Theory Rel. Fields, **118**, 65-114, (2000),

[CM04] O. COURONNÉ and R. J. MESSIKH, Surface order large deviations for 2D FK-percolation and Potts models, Stoc. Proc. Appl. 113 (2004) 81-99.

[DZ98] A. DEMBO and O. ZEITOUNI, *Large Deviations Techniques and Applications,* Second edition, Springer, New York (1998).

[DRS14] A. DREWITZ, B. RÁTH, and A. SAPOZHNIKOV, On chemical distances and shape theorems in percolation models with long-range correlations. J. Math. Phys. **55** (2014), no. 8, 083307, 30 pp.

[FP16] A. FRIBERGH AND S. POPOV Biased random walk on the interlacement set, arXiv:1610.02979.

[GM04] O. GARET and R. MARCHAND, Asymptotic shape for the chemical distance and first-passage percolation on the infinite Bernoulli cluster, ESAIM Probab. Stat. 8 (2004), 169-199.

[GRS14] N. GIORGIOU, F. RASSOUL-AGHA and T. SEPPÄLÄINEN, Variational formulas and cocycle solutions for directed polymer and percolation model, *Preprint*, (2014)

[GRSY13] N. GIORGIOU, F. RASSOUL-AGHA, T. SEPPÄLÄINEN and A. YILMAZ, Ratios of partition functions for the log-gamma polymer, *Ann. Probab*, to appear, (2013)

[GdH98] A. GREVEN and F. DEN HOLLANDER, Large deviations for a random walk in a random environment, *Ann. Prob.*, **22**, 1381-1428, (1998)

[G99] G. R. GRIMMETT, Percolation, Springer-Verlag, 2006.

[G06] G. R. GRIMMETT, The random-cluster model, Springer-Verlag, 2006.

[GM90] G. GRIMMETT and J. MARSTRAND, The supercritical phase of percolation is well behaved, Proc. R. Soc. Lond. Ser. A 430 (1990) 439-457.

[K75] H. KESTEN Sums of stationary sequences can not grow slower than linearly, Proc. AMS, **49**, 205-211, (1975)

[KV86] C. KIPNIS and S.R.S. VARADHAN, Limit theorem for additive functionals of reversible Markov chains and application to simple exclusions, *Comm. Math. Phys.* **104**, 1-19, (1986)

[KRV06] E. KOSYGINA, F. REZAKHANLOU and S. R. S. VARADHAN Stochastic homogenization of Hamilton-Jacobi-Bellmann equations, *Comm. Pure Appl. Math.*, **59**, 1489-1521, (2006)

[K85] S. M. KOZLOV, The averaging effect and walks in inhomogeneous environments, *Uspekhi Mat Nayuk*, (Russian math surveys), **40**, 73-145, (1985)

[K12] N. KUBOTA, Large deviations for simple random walks on supercritical percolation clusters, Kodai Mathematical Journal **35**, Number 3, 560-575, (2012)

[LS86] J. L. LEBOWITZ and H. SALEUR. Percolation in strongly correlated systems. Phys. A, 138: pp. 194-205 (1986).

[LSS97] T. LIGGETT, R. SCHONMANN and A. STACEY, Domination by product measures, Ann. Probab. **25** (1997) 71-95.

[M12] J-C. MOURRAT Lyapunov exponents, shape theorems and large deviations for random walks in random potential, *ALEA Lat. Am. J. Probab. Math. Stat.* **9**, 165-211, (2012)

[MP07] P. MATHEIU and A. PIATNITSKI, Quenched invariance principle for random walks on percolation clusters , Proceedings of the Royal Society A., 463, 2287-2307, (2007)

[PV81] G. C. PAPANICOLAOU and S. R. S. VARADHAN, Boundary value problems with rapidly os- cillating random coefficients, In Random fields, Vol. I, II (Esztergom, 1979), volume 27 of Colloq. Math. Soc. Janos Bolyai, pages 835ÃČÂŘ873. North-Holland, Amsterdam, 1981

[P96] A. PISZTORA, Surface order large deviations for Ising, Potts and percolation models, Probab. Theory Relat. Fields **104** (1996) 427-466.

[PRS15] E. B. PROCACCIA, R. ROSENTHAL and A. SAPOZHNIKOV, Quenched invariance principle for simple random walk on clusters in correlated percolation models, Probab. Theory Relat. Fields DOI 10.1007/s00440-015-0668-y.

[R70] R. T. ROCKAFELLAR, Convex analysis, Princeton University Press, Princeton, N. J., 1997.

[R15] P.-F. RODRIGUEZ, A 0-1 law for the massive Gaussian free field, preprint, arXiv:1505.08169.

[R06] J. ROSENBLUTH Quenched large deviations for multidimensional random walks in a random environment: a variational formula, *PhD thesis*, NYU, arxiv:0804.1444v1

[RS11] F. RASSOUL-AGHA and T. SEPPÄLÄINEN, Process-level quenched large deviations for random walk in a random environment, *Ann. Inst. H. Poincaré Prob. Statist.*, **47**, 214-242, (2011)

[RSY13] F. RASSOUL-AGHA, T. SEPPÄLÄINEN and A. YILMAZ, Quenched free energy and large deviations for random walk in random potential, *Comm. Pure and Appl. Math*, **66**, 202-244, (2013)

[RSY14] F. RASSOUL-AGHA, T. SEPPÄLÄINEN and A. YILMAZ, Variational formulas and disorder regimes of random walks in random potential, *Preprint*

[RS13] P.-F. RODRIGUEZ AND A.-S. SZNITMAN Phase transition and level-set percolation for the Gaussian free field, Commun. Math. Phys. 320, 571-601 (2013)

[S07] S. SHEFIELD Gaussian free fields for mathematicians, Probab. Theory and Related Fields 139 (2007) 521-541.

[SS04] V. SIDORAVICIUS and A. S. SZNITMAN, Quenched invariance principles for walks on clusters of percolation or among random conductances, Probability theory and related fields., 129, 219-244, (2004)

[S94] A. S. SZNITMAN, Shape theorem Lyapunov exponents and large deviations for Brownian motion in a Poissonian potential, *Comm. Pure. Appl. Math.*, **47**, 1655-1688, (1994)

[S03] A. S. SZNITMAN, On the anisotropic walk on the supercritical percolation cluster, *Comm. Math. Phys.*, 240, Issue 1-2, 123-148, (2003)

[S10] A. S. SZNITMAN, Vacant set of random interlacements and percolation, Ann. Math., 171 (2), 2039-2087, (2010).

[T09] A. TEIXEIRA, Interlacement percolation on transient weighted graphs, Elec. J. Probab. 14 (2009), no. 54, 1604-1628.

[T09'] A. TEIXEIRA, On the uniqueness of the infinite cluster of the vacant set of random interlacements, Adv. Appl. Prob. 19 (2009), 454-466.

[TW11] A. TEIXEIRA AND D. WINDISCH, On the fragmentation of a torus by random walk. Comm. Pure Appl. Math. 64 (12), 1599-Ş1646.

[V03] S. R. S. VARADHAN, large deviations for random walk in random environment, Comm. Pure Appl. Math, **56**, Issue 8, 1222-1245, (2003)

[Y08] A. YILMAZ, Quenched large deviations for random walk in random environment, *Comm. Pure Appl. Math*, **62**, Issue 8, 1033- 1075, (2009)

[Z98] M. ZERNER, Lyapunov exponents and quenched large deviations for multidimensional random walks in random environment, *Ann. Prob.*, **26**, No. 4, 1446-1476, (1998)