Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2019, Vol. 55, No. 2, 862–881
https://doi.org/10.1214/18-AIHP901

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

The speed of biased random walk among random conductances

Noam Berger[a], Nina Gantert[a] and Jan Nagel[b]

[a]*Technische Universität München, Fakultät für Mathematik, Boltzmannstraße 3, 85748 Garching bei München, Germany.*
E-mail: noam.berger@tum.de; nina.gantert@ma.tum.de
[b]*Eindhoven University of Technology, Department of Mathematics and Computer Science, 5600 MB Eindhoven, Netherlands.*
E-mail: j.h.nagel@tue.nl

Received 27 April 2017; revised 13 March 2018; accepted 23 March 2018

Abstract. We consider biased random walk among iid, uniformly elliptic conductances on $\mathbb{Z}^d$, and investigate the monotonicity of the velocity as a function of the bias. It is not hard to see that if the bias is large enough, the velocity is increasing as a function of the bias. Our main result is that if the disorder is small, i.e. all the conductances are close enough to each other, the velocity is always strictly increasing as a function of the bias, see Theorem 1.1. A crucial ingredient of the proof is a formula for the derivative of the velocity, which can be written as a covariance, see Theorem 1.3: it follows along the lines of the proof of the Einstein relation in (*Ann. Probab.* **45** (4) (2017) 2533–2567). On the other hand, we give a counterexample showing that for iid, uniformly elliptic conductances, the velocity is not always increasing as a function of the bias. More precisely, if $d = 2$ and if the conductances take the values 1 (with probability p) and κ (with probability $1 - p$) and p is close enough to 1 and κ small enough, the velocity is *not* increasing as a function of the bias, see Theorem 1.2.

Résumé. Nous étudions des marches aléatoires biaisées dans un milieu aléatoire donné par des poids iid sur les arêtes de $\mathbb{Z}^d$. Les poids sont bornés au-dessus et ils ont une borne inférieure qui est strictement positive. Nous nous intéressons pour la vitesse de la marche en fonction du bias. Un argument connu donne que, pour des biais suffisamment grands, la vitesse est une fonction croissante du biais. Notre résultat principal dit que si le désordre est petit, ce qui veut dire que les poids sont proches les uns aux autres, la vitesse est une fonction croissante du bias, voir Théorème 1.1. Un ingrédient crucial de la preuve est une formule pour la dérivée de la vitesse : cette dérivée peut etre écrit comme une covariance, voir Théorème 1.3. La preuve de Théorème 1.3 suis les arguments de la preuve de la relation d'Einstein dans (*Ann. Probab.* **45** (4) (2017) 2533–2567). Par contre, nous donnons un exemple montrant que pour des poids iid prenant les valeurs 1 (avec probabilité p) et κ (avec probabilité $1 - p$), si p est suffisamment proche de 1 et κ est suffisamment petit, la vitesse n'est *pas* une fonction croissante du bias, voir Théorème 1.2.

MSC: 60K37; 60G42; 60J10

Keywords: Random walk in random environment; Random conductances; Effective velocity; Regeneration times

1. Introduction

As a model for transport in an inhomogeneous medium, one may consider a biased random walk on a supercritical percolation cluster. The model goes back, to our best knowledge, to Mustansir Barma and Deepak Dhar, see [2] and [8]. They conjecured the following picture for the velocity (in the direction of the bias) as a function of the bias. The velocity is increasing for small values of the bias, then it is decreasing to 0 and remains 0 for large values of the bias, see Figure 2 below. Here, the zero velocity regime is due to "traps" in the environment which slow down the random walk. It was proved by [19] and by [5] that the velocity is indeed zero if the bias is large enough, while it is strictly positive for small values of the bias. Later, Alexander Fribergh and Alan Hammond were able to show that there is a sharp transition, i.e. there is a critical value of the bias such that the velocity is zero if the bias is larger, and strictly positive if the bias is smaller than the critical value, see [11].

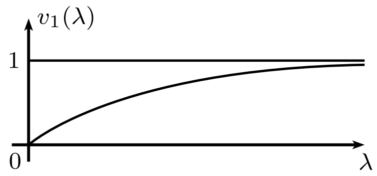

Fig. 1. Speed of biased simple random walk.

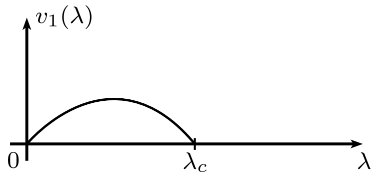

Fig. 2. Conjectured speed of biased random walk on percolation clusters.

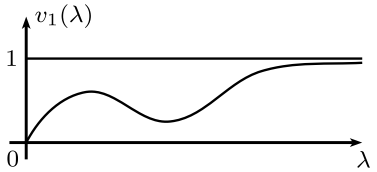

Fig. 3. Conjectured speed of biased random walk under the assumptions of Theorem 1.2.

The velocity of biased random walk among iid, uniformly elliptic conductances is always strictly positive, this was proved by Lian Shen in [18]. A criterion for ballisticity in the elliptic, but not uniformly elliptic case can be found in [10]. It is interesting to ask about monotonicity in the uniformly elliptic case. In the following, $v_1(\lambda)$ denotes the component of the velocity in the direction of the bias, precise definitions are below. In the homogeneous medium (i.e. if the conductances are constant), the velocity can be computed and the picture is as in Figure 1. For the biased random walk on a (supercritical) percolation cluster, the conjectured picture is as in Figure 2. Now, in our case of iid, uniformly elliptic conductances, the picture should be "in between" the other two cases. If the conductances are close enough to each other, we show that the speed is increasing, hence the picture is as in Figure 1. Under the assumptions of Theorem 1.2, we show that the speed is not increasing and Figure 3 is the simplest picture which agrees with our results. However, we only prove parts of this picture: we know that for $\lambda \to \infty$, the velocity is increasing and goes to 1, see Fact 2 below, and we show that the velocity is not increasing for all values of the bias, see Theorem 1.2.

Finally, let us mention some results for biased random walks on supercritical Galton–Watson trees with a bias pointing away from the root. This model can be seen as a "toy model" for the percolation case, when the lattice is replaced by a tree. For biased random walks on (supercritical) Galton–Watson trees with leaves, the velocity shows the same regimes as for biased random walks on percolation clusters: it is zero if the bias is larger than a critical value, while it is strictly positive if the bias is less (or equal) than the critical value. This transition was proved by [16] and the critical value has an explicit description, see [16]. In particular, if the tree has leaves, the velocity can not be an increasing function of the bias. For biased random walks on supercritical Galton–Watson trees without leaves the velocity is conjectured to be increasing, but despite recent progress, see [1,4], this conjecture is still open.

Let us now give more precise statements and a description of our results. For two neighboring vertices x and y in $\mathbb{Z}^d$ with $d \geq 2$, assign to the edge between x and y a nonnegative *conductance* $\omega(x, y)$. The random walk among the conductances ω starting at x_0 and with bias $\lambda \geq 0$ (in direction $e_1 = (1, 0, 0, \dots, 0)$) is then the Markov chain $(X_n)_{n\geq 0}$ with law $P^{x_0}_{\omega,\lambda}$, defined by the transition probabilities

$$P^{x_0}_{\omega,\lambda}(X_{n+1} = y | X_n = x) = \frac{\omega(x, y)e^{\lambda(y-x)\cdot e_1}}{\sum_{z\sim x} \omega(x, z)e^{\lambda(z-x)\cdot e_1}}$$

for $x \sim y$. (Here we write $x \sim y$ if x, y are neigboring vertices, and we write $w \cdot z$ for the scalar product of two vectors $w, z \in \mathbb{R}^d$.) The corresponding expectation is written as $E^{x_0}_{\omega,\lambda}$. The Markov chain $(X_n)_{n\geq 0}$ is reversible with respect to the measure

$$\pi(x) = \sum_{z\sim x} \omega(x, z) e^{\lambda(x+z)\cdot e_1}. \tag{1.1}$$

When the collection of conductances ω is random with law P, we call $(X_n)_{n\geq 0}$ *random walk among random conductances* and $P^{x_0}_{\omega,\lambda}$ the *quenched law*. $\mathbb{P}^{x_0}_{\lambda} = \int P^{x_0}_{\omega,\lambda}(\cdot) P(d\omega)$ is called the *annealed law* and we write $\mathbb{E}^{x_0}_{\lambda}$ for the corresponding expectation. If $x_0 = 0$ we omit the superscripts. In this paper we study properties of the limiting velocity

$$v(\lambda) = \lim_{n\to\infty} \frac{X_n}{n}. \tag{1.2}$$

Frequently, we focus on the speed in direction e_1 and set $v_1(\lambda) = v(\lambda) \cdot e_1$. In particular, we are interested in the monotonicty of v_1 as a function of the bias λ. Although increasing λ increases the local drift to the right at every point, it is not clear at all that this results in a higher effective velocity. As mentioned above, this conclusion is known to be false for a biased random walk on a percolation cluster, which corresponds to conductances $\omega(x, y) \in \{0, 1\}$. As shown by [11], the speed is positive for λ smaller than some critical value $\lambda_c > 0$, but increasing the bias further will give zero speed. If we assume the conductances to be uniformly elliptic, that is, there exists a $\delta \in (0, 1)$ such that

$$1 - \delta \leq \omega(x, y) \leq 1 + \delta, \tag{1.3}$$

then [18] showed that the limit in (1.2) exists $\mathbb{P}_\lambda$ almost surely, does not depend on ω, and there is no zero speed regime: $v_1(\lambda) > 0$ for all $\lambda > 0$. From now on, we assume

Assumption (A). The conductances are iid and uniformly elliptic, i.e. they satisfy (1.3).

Note that (1.3) is equivalent to the usual uniform ellipticity saying that the conductances are bounded above and bounded away from 0: we may multiply all the conductances by a constant factor, resulting in the same transition probabilities.

Fact 1. $\lim_{\lambda\to\infty} v_1(\lambda) = 1$.

Fact 2. *There exists a* $\lambda_c = \lambda_c(\delta)$ *such that* v_1 *is strictly increasing on* $[\lambda_c, \infty)$.

Fact 1 follows from a coupling with a random walk in a homogeneous environment, as

$$P_{\omega,\lambda}(X_{n+1} = x + e_1 | X_n = x) \geq \frac{e^\lambda}{(2d-1)\frac{1+\delta}{1-\delta} + e^\lambda}, \tag{1.4}$$

which goes to 1 as $\lambda \to \infty$. Fact 2 was proven by [4] for the biased random walk on a Galton–Watson tree without leaves (where an upper bound for λ_c can be explicitly computed), the same arguments yield the analogous result for the conductance model, when the conductances are bounded away from 0 and ∞. A sketch of the proof will be given in Section 2. We remark that $\lambda_c(\delta)$ may be chosen decreasing in δ.

Our first main result shows that in the low disorder regime, when δ is close to 0, v_1 is increasing on $[0, \infty)$. That is, in the low disorder regime, Fact 2 holds with $\lambda_c = 0$.

Theorem 1.1. *Assume* (A). *There exists a* $\delta_0 \in (0, 1)$, *such that if* $1 - \delta_0 \leq \omega(x, y) \leq 1 + \delta_0$ *whenever* $x \sim y$, *then* v_1 *is strictly increasing*.

On the other hand, outside the low disorder regime, there is in general no monotonicity, in particular, uniform ellipticity of the conductances does not imply monotonicity of the speed.

Theorem 1.2. *Assume* (A) *and* $d = 2$. *Define the environment law by*

$$P\big(\omega(0,e) = 1\big) = p = 1 - P\big(\omega(0,e) = \kappa\big)$$

for $p \in (0,1)$ *and* $\kappa > 0$. *Then, for* p *close enough to* 1 *and* κ *close enough to* 0, *there exist* $\lambda_1 < \lambda_2$ *such that*

$$v_1(\lambda_1) > v_1(\lambda_2).$$

To prove Theorem 1.1, we show that the derivative of the speed is strictly positive, where the derivative can be expressed as the covariance of two processes. For this, we define

$$M_n = X_n - \sum_{k=0}^{n-1} E_{\omega,\lambda}^{X_k}[X_1 - X_0], \tag{1.5}$$

$$N_n = X_n - n v(\lambda). \tag{1.6}$$

We show in Proposition 3.1 below that under $\mathbb{P}_\lambda$, the $2d$-dimensional process $\frac{1}{\sqrt{n}}(M_n, N_n)$ converges in distribution to a Gaussian limit (M, N).

Theorem 1.3. *Assume* (A). *For any* $\lambda > 0$, v *is differentiable at* λ *with*

$$v'(\lambda) = \mathrm{Cov}_\lambda(M, N)e_1.$$

Remark 1.1. The statement in Theorem 1.3 is true for $\lambda = 0$ as well – this is the Einstein relation proved in [12]. In particular, $\lambda \to v_1(\lambda)$ is a continuous function. The continuity of v_1 may seem obvious, but to our best knowledge, it has not been proved for a biased random walk on a percolation cluster.

2. A general coupling

After a suitable enlargement of our probability space, let $U_0, U_1, \ldots$ be a sequence of independent random variables with a uniform distribution on $[0,1]$, independent of ω. Let us denote the joint law of the U_k and ω by $\mathbb{P}$, with expectation $\mathbb{E}$. We will construct a coupling of quenched laws for different environments and different values of the bias, letting U_k determine the movement at time k. Given an environment ω and $\lambda \geq 0$, define

$$p_{\omega,\lambda}(x,e) = P_{\omega,\lambda}(X_1 = x + e | X_0 = x)$$

and, with $e_k = -e_{2d+1-k}$ for $d+1 \leq k \leq 2d$, let $q_{\omega,\lambda}(x,0) = 0$ and for $1 \leq k \leq 2d$,

$$q_{\omega,\lambda}(x,k) = \sum_{j=1}^{k} p_{\omega,\lambda}(x, e_j).$$

Now, given two environments ω_1 and ω_2 and biases λ_1 and λ_2 we can define processes $X_n^{(\omega_1,\lambda_1)}$ and $X_n^{(\omega_2,\lambda_2)}$ by setting

$$X_{n+1}^{(\omega_i,\lambda_i)} - X_n^{(\omega_i,\lambda_i)} = e_k \quad \text{iff} \quad q_{\omega_i,\lambda_i}(X_n, k-1) < U_n \leq q_{\omega_i,\lambda_i}(X_n, k)$$

for $i = 1, 2$. Then the marginal of $(X_n^{(\omega_i,\lambda_i)})_n$ is the original quenched law P_{ω_i,λ_i}. In the one-dimensional case this coupling also shows the monotonicity of the speed for any ellipticity constant, since then $\lambda_1 \leq \lambda_2$ implies $X_n^{(\omega,\lambda_1)} \leq X_n^{(\omega,\lambda_2)}$. To give a short justification of Fact 2, we additionally introduce for $\lambda_s > 0$ the one-dimensional process

$$Y_n = \sum_{k=0}^{n-1}\left(2 \cdot \mathbb{1}\left\{U_k \leq \frac{e^{\lambda_s}}{e^{\lambda_s} + (2d-1)\beta}\right\} - 1\right),$$

where $\beta = \frac{1+\delta}{1-\delta}$. Assume $\lambda_s > \log\beta + \log(2d-1)$, then Y_n is a simple random walk with drift to the right. From the lower bound (1.4), we see that if Y_n moves to the right and $\lambda_s < \lambda_i$, then $X_n^{(\omega_i,\lambda_i)}$ moves to the right. This allows us to consider so-called super-regeneration times τ_k, $k \geq 1$ (introduced by [3]) where τ_1 is the infimum over all times $n \geq 1$ with

$$\max_{k<n-1} Y_k < Y_{n-1} < Y_n < \min_{k>n} Y_k,$$

and inductively $\tau_{n+1} = \tau_1 \circ \theta_{\tau_n} + \tau_n$ (here θ_k denotes the time shift, i.e. $\theta_k(Y_n)_{n\geq 0} = (Y_{k+n})_{n\geq 0}$). Since the increments of Y_n are a lower bound for the increments of $X_n^{(\omega_i,\lambda_i)}$ in direction e_1, τ_1 is a regeneration time for the process $X_n = X_n^{(\omega_i,\lambda_i)}$, provided $\lambda_i > \lambda_s$. More precisely,

$$\max_{k<\tau_n-1} X_k \cdot e_1 < X_{\tau_n-1} \cdot e_1 < X_{\tau_n} \cdot e_1 < \min_{k>\tau_n} X_k \cdot e_1.$$

Unlike in [4], we require an additional step to the right in order to decouple the environment seen by the random walker. By classical arguments, the sequence $(X_{\tau_k}^{(\omega,\lambda)} - X_{\tau_{k-1}}^{(\omega,\lambda)}, \tau_k - \tau_{k-1})_{k\geq 2}$ is an iid sequence under $\mathbb{P}$, and the marginal is equal to the distribution of (X_{τ_1}, τ_1), conditioned on the event $R = \{Y_n > 0 \text{ for all } n \geq 1\}$. Moreover,

$$v(\lambda) = \frac{\mathbb{E}[X_{\tau_1}^{(\omega,\lambda)} | R]}{\mathbb{E}[\tau_1 | R]}$$

for any $\lambda > \lambda_s$. Fact 2 follows then if we can show for λ_s large enough and $\lambda > \lambda_s$,

$$\mathbb{E}\big[\big(X_{\tau_1}^{(\omega,\lambda+\varepsilon)} - X_{\tau_1}^{(\omega,\lambda)}\big) \cdot e_1 | R\big] > 0 \tag{2.1}$$

for any $\varepsilon > 0$. Following the arguments of [4], this is implied by the following observations:

- When Y_n moves to the right, both $X_n^{(\omega,\lambda)}$ and $X_n^{(\omega,\lambda+\varepsilon)}$ move to the right.
- When Y_n moves to the left for the first time, then

  $$\big(X_n^{(\omega,\lambda+\varepsilon)} - X_n^{(\omega,\lambda)}\big) \cdot e_1 \geq 0$$

  and, given that Y_n moves to the left for the first time at time n, we have with positive probability

  $$\big(X_n^{(\omega,\lambda+\varepsilon)} - X_n^{(\omega,\lambda)}\big) \cdot e_1 > 0. \tag{2.2}$$

  In fact, given that $X_n^{(\omega,\lambda+\varepsilon)}$ and $X_n^{(\omega,\lambda)}$ are at x at time $n-1$ and decouple at time n for the first time, (2.2) occurs with probability

  $$\frac{q_{\omega,\lambda+\varepsilon}(x,1) - q_{\omega,\lambda}(x,1) + q_{\omega,\lambda+\varepsilon}(x,2d-1) - q_{\omega,\lambda}(x,2d-1)}{\sum_{k=1}^{2d-1} |q_{\omega,\lambda+\varepsilon}(x,k) - q_{\omega,\lambda}(x,k)|} \geq \frac{c(\delta)}{2d-1}. \tag{2.3}$$

  To see the lower bound in (2.3), note that $q_{\omega,\lambda+\varepsilon}(x,2d-1) - q_{\omega,\lambda}(x,2d-1) \geq 0$. Further, recalling (1.1), a small calculation shows that $\pi(\lambda)\pi(\lambda+\varepsilon)(q_{\omega,\lambda+\varepsilon}(x,1) - q_{\omega,\lambda}(x,1)) \geq c_1(\delta)e^{\lambda}\varepsilon$, and $\pi(\lambda)\pi(\lambda+\varepsilon)|q_{\omega,\lambda+\varepsilon}(x,k) - q_{\omega,\lambda}(x,k)| \leq c_2(\delta)e^{\lambda}\varepsilon$ for all k, resulting in (2.3).
- When until time τ_1 the process Y_n took k steps to the left, the increments of $X_n^{(\omega,\lambda)}$ and $X_n^{(\omega,\lambda+\varepsilon)}$ could differ at most k times.
- When until time n the increments of $X_n^{(\omega,\lambda)}$ and $X_n^{(\omega,\lambda+\varepsilon)}$ were different exactly k times, then

  $$\big(X_{\tau_1}^{(\omega,\lambda+\varepsilon)} - X_{\tau_1}^{(\omega,\lambda)}\big) \cdot e_1 > -2(k-1).$$

- Let D_k be the event that until time τ_1, Y_n did k steps to the left and for some $n \leq \tau_1$, $X_n^{(\omega,\lambda+\varepsilon)} - X_n^{(\omega,\lambda)} \neq 0$. Then

  $$\mathbb{E}\big[\big(X_{\tau_1}^{(\omega,\lambda+\varepsilon)} - X_{\tau_1}^{(\omega,\lambda)}\big) \cdot e_1 | R\big] \geq (2d-1)^{-1}\mathbb{P}(D_1|R) - \sum_{k\geq 2} 2(k-1)\mathbb{P}(D_k|R). \tag{2.4}$$

For λ_s large enough, the right hand side of (2.4) is positive, which follows analogously to the proof in [4] of positivity of display (4.1) therein.

3. Differentiating the speed

Theorem 1.3 is a consequence of the two following results. For simplicity, we will omit integer parts.

Theorem 3.1. *Let* $\lambda_0 > 0$, $\alpha > 1$ *and* $t_\lambda = \alpha \cdot (\lambda - \lambda_0)^{-2}$, *then*

$$\lim_{\lambda \to \lambda_0} \frac{\frac{1}{t_\lambda}\mathbb{E}_\lambda[X_{t_\lambda}] - v(\lambda_0)}{\lambda - \lambda_0} = \mathrm{Cov}_{\lambda_0}(M, N) \cdot e_1.$$

Theorem 3.2. *Let* t_λ *be as in Theorem* 3.1. *There exists a* $C > 0$, *such that for any* $\alpha > 1$,

$$\limsup_{\lambda \to \lambda_0} \left| \frac{\frac{1}{t_\lambda}\mathbb{E}_\lambda[X_{t_\lambda}] - v(\lambda)}{\lambda - \lambda_0} \right| \le \frac{C}{\sqrt{\alpha}}.$$

3.1. *Regeneration times*

The proof of Theorem 3.1 and Theorem 3.2 relies on a regeneration structure for the process $(X_n)_n$, which decomposes the trajectory into 1-dependent increments with good moment bounds. For $h \in \mathbb{R}$, we let

$$\mathcal{H}_h = \{x \in \mathbb{Z}^d | x \cdot e_1 = \lfloor h \rfloor\}$$

denote the hyperplane with first coordinate $\lfloor h \rfloor$ and

$$T_h = \inf\{n \ge 0 | X_n \in \mathcal{H}_h\}$$

be the first hitting time of $\mathcal{H}_h$. The regeneration times τ_k, $k \ge 1$ are then hitting times $T_{mL/\lambda}$, after which the random walk never visits $\mathcal{H}_{(m-1)L/\lambda}$ again and the displacement $X_{T_{mL/\lambda}} - X_{T_{(m-1)L/\lambda}}$ can be decoupled from the environment in $\{x \in \mathbb{Z}^d | x \cdot e_1 \le \lfloor h \rfloor\}$. The detailed construction of the sequence $(\tau_k)_k$ can be found in [12], for the sake of brevity we only summarize here the consequences in the following lemma. We remark that the moment bounds are stated in [12] only for $\lambda \in (0, \lambda_u)$ for some small $\lambda_u > 0$, but the proof works actually for any bounded, positive λ.

Remark 3.1. Note that the $(\tau_k)_k$ are not the same as the super-regeneration times in Section 2 (which were also denoted by $(\tau_k)_k$) but in order to be consistent with [3] and [12], we keep this notation.

Lemma 3.1. *Under* $\mathbb{P}_\lambda$, *the sequence*

$$\big((X_{k+1} - X_k)_{\tau_n \le k < \tau_{n+1}}, \tau_{n+1} - \tau_n\big)_{n \ge 1}$$

is a stationary 1*-dependent sequence. Moreover, for any* $\lambda_1 > 0$ *there are constants* $c, C > 0$, *such that for all* $\lambda \in (0, \lambda_1]$ *we have*

$$\mathbb{E}_\lambda\big[\exp\big(c\lambda^2 \tau_1\big)\big] \le C, \qquad \mathbb{E}_\lambda\big[\exp\big(c\lambda^2(\tau_2 - \tau_1)\big)\big] \le C \tag{3.1}$$

and

$$\mathbb{E}_\lambda\big[\exp\big(c\lambda \|X_{\tau_1}\|\big)\big] \le C, \qquad \mathbb{E}_\lambda\big[\exp\big(c\lambda \|X_{\tau_2} - X_{\tau_1}\|\big)\big] \le C.$$

We also have a lower bound for the inter-regeneration time (see (21) in [12]), where for any $\lambda_1 > 0$ there is a constant $c > 0$, such that

$$\mathbb{E}_\lambda\big[\lambda^2(\tau_2 - \tau_1)\big] \ge c \tag{3.2}$$

for all $\lambda \in (0, \lambda_1]$. If (1.3) is satisfied with $\delta \le \frac{1}{2}$, c and C in Lemma 3.1 and in (3.2) can be chosen only depending on the dimension. As a consequence of Lemma 3.1 and the law of large numbers, we get the following expression for the speed,

$$v(\lambda) = \frac{\mathbb{E}_\lambda[X_{\tau_2} - X_{\tau_1}]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]}, \tag{3.3}$$

see Corollary 16 in [12].

Using the exponential moment estimates on the regeneration times, it follows that in order to study the convergence in distribution of $\frac{1}{\sqrt{n}}(M_n, N_n)$, it suffices to consider

$$\frac{1}{\sqrt{\tau_n}}(M_{\tau_n}, N_{\tau_n}).$$

To this subsequence, we may apply the functional central limit theorem for sums of 1-dependent random variables, see [6] to obtain the following result.

Proposition 3.1. *For any $\lambda > 0$, the process $(\frac{1}{\sqrt{n}}(M_{\lfloor tn \rfloor}, N_{\lfloor tn \rfloor}); t \ge 0)$ converges in distribution under $\mathbb{P}_\lambda$ to a $2d$-dimensional Brownian motion $(\widehat{M}_t, \widehat{N}_t)$. We write M for $\widehat{M}_1$ and N for $\widehat{N}_1$.*

Lemma 3.2. *For any $p \in \mathbb{N}$ and $\lambda_1 > 0$ there exists a $C_p > 0$ depending only on p, λ_1, the dimension d, and the ellipticity constant δ, such that for any $0 < \lambda < \lambda_1$,*

$$\mathbb{E}_\lambda\Big[\max_{0 \le k \le n/\lambda^2} \|\lambda X_k\|^p\Big] \le C_p n^p.$$

Proof. The lemma follows from the proof of Lemma 8 in [12], noting that the constant C_p there can be chosen depending only on p, an upper bound for λ, the dimension d, and the ellipticity constant δ. □

3.2. *Proof of Theorem* 3.1

The arguments in this section are inspired by [15] where a weak form of the Einstein relation was proved for a large class of models. Let us abbreviate $\bar{\lambda} = \lambda - \lambda_0$ and begin by writing, with $t = t_\lambda = \alpha/\bar{\lambda}^2$,

$$\frac{\frac{1}{t}\mathbb{E}_\lambda[X_t] - v(\lambda_0)}{\lambda - \lambda_0} = \mathbb{E}_\lambda\left[\frac{\bar{\lambda}}{\alpha}\big(X_t - t \cdot v(\lambda_0)\big)\right] = \mathbb{E}_{\lambda_0}\left[\frac{\bar{\lambda}}{\alpha}\big(X_t - t \cdot v(\lambda_0)\big)\frac{dP_{\omega,\lambda}}{dP_{\omega,\lambda_0}}(X_s; 0 \le s \le t)\right] \tag{3.4}$$

as an expectation with respect to the reference measure $\mathbb{P}_{\lambda_0}$. For a nearest-neighbor path $(x_1, \dots, x_m)$, we have

$$\frac{dP_{\omega,\lambda}}{dP_{\omega,\lambda_0}}(x_1, \dots, x_m) = \prod_{k=1}^m \frac{p_{\omega,\lambda}(x_{k-1}, x_k - x_{k-1})}{p_{\omega,\lambda_0}(x_{k-1}, x_k - x_{k-1})} = \prod_{k=1}^m e^{\bar{\lambda}(x_k - x_{k-1})\cdot e_1} \frac{\sum_{|e|=1} e^{\lambda_0 e\cdot e_1}\omega(x_{k-1}, x_{k-1}+e)}{\sum_{|e|=1} e^{\lambda e\cdot e_1}\omega(x_{k-1}, x_{k-1}+e)}.$$

Now write in the denominator $e^{\lambda e\cdot e_1} = e^{\bar{\lambda} e\cdot e_1} e^{\lambda_0 e\cdot e_1}$ and expand the first exponential $e^z = 1 + z + z^2/2 + r_1(z)$ with $|r_1(z)| \le |z|^3$ for $|z| \le 1$ to get

$$\begin{aligned}&\frac{dP_{\omega,\lambda}}{dP_{\omega,\lambda_0}}(x_1, \dots, x_m)\\ &\quad = \exp\left\{\bar{\lambda} x_m \cdot e_1 - \sum_{k=1}^m \log\left(1 + \bar{\lambda} d_{\omega,\lambda_0}(x_{k-1}) + \frac{1}{2}\bar{\lambda}^2 d^{(2)}_{\omega,\lambda_0}(x_{k-1}) + r_1(\bar{\lambda})\right)\right\},\end{aligned}$$

where we wrote

$$d_{\omega,\lambda_0}(x)=\frac{\sum_{|e|=1}\omega(x,x+e)e^{\lambda_0 e\cdot e_1}e\cdot e_1}{\sum_{|e|=1}\omega(x,x+e)e^{\lambda_0 e\cdot e_1}}=E^x_{\omega,\lambda_0}\big[(X_1-X_0)\cdot e_1\big]$$

for the local drift in direction e_1 and

$$d^{(2)}_{\omega,\lambda_0}(x)=\frac{\sum_{|e|=1}\omega(x,x+e)e^{\lambda_0 e\cdot e_1}(e\cdot e_1)^2}{\sum_{|e|=1}\omega(x,x+e)e^{\lambda_0 e\cdot e_1}}=E^x_{\omega,\lambda_0}\big[\big((X_1-X_0)\cdot e_1\big)^2\big]$$

for the expected squared displacement. Expanding the logarithm as $\log(1+z)=z-z^2/2+r_2(z)$ with $|r_2(z)|\le|z|^3$ for $|z|\le 1/2$, we obtain

$$\exp\Bigg\{\bar\lambda x_m\cdot e_1-\sum_{k=1}^m\bigg(\bar\lambda d_{\omega,\lambda_0}(x_{k-1})+\frac{\bar\lambda^2}{2}\big(d^{(2)}_{\omega,\lambda_0}(x_{k-1})-d_{\omega,\lambda_0}(x_{k-1})^2\big)+h(\bar\lambda)\bigg)\Bigg\},$$

where the function h satisfies $|h(z)|\le c|z|^3$ if $|z|\le 1/2$. If we set now $m=t=\alpha/\bar\lambda^2$, this yields

$$\begin{aligned}G_{\omega,\lambda_0}(\bar\lambda,t)&:=\frac{dP_{\omega,\lambda}}{dP_{\omega,\lambda_0}}(X_k;0\le k\le t)\\&=\exp\Bigg\{\bar\lambda\Bigg(X_{\alpha/\bar\lambda^2}\cdot e_1-\sum_{k=1}^{\alpha/\bar\lambda^2}d_{\omega,\lambda_0}(X_{k-1})\Bigg)-\frac{\bar\lambda^2}{2}\sum_{k=1}^{\alpha/\bar\lambda^2}\big(d^{(2)}_{\omega,\lambda_0}(X_{k-1})-d_{\omega,\lambda_0}(X_{k-1})^2\big)+o(1)\Bigg\}.\end{aligned}\tag{3.5}$$

By Proposition 3.1, $\bar\lambda(X_{\alpha/\bar\lambda^2}\cdot e_1-\sum_{k=1}^{\alpha/\bar\lambda^2}d_{\omega,\lambda_0}(X_{k-1}))$ converges in distribution to $\widehat{M}_\alpha\cdot e_1$. To infer the convergence of the complete expression for the density and to obtain convergence of the expectations in (3.4), we next show L^p-boundedness of the density for $\bar\lambda$ small enough.

Recall $G_{\omega,\lambda_0}(\bar\lambda,t)$ in (3.5), and let $p\ge 1$. Then

$$\begin{aligned}p\log G_{\omega,\lambda_0}(\bar\lambda,t)&=p\bar\lambda X_t\cdot e_1-p\sum_{k=1}^t\log\bigg(\frac{\sum_{|e|=1}\omega(X_{k-1},X_{k-1}+e)e^{\lambda e\cdot e_1}}{\sum_{|e|=1}\omega(X_{k-1},X_{k-1}+e)e^{\lambda_0 e\cdot e_1}}\bigg)\\&=p\bar\lambda X_t\cdot e_1-\sum_{k=1}^t\log\bigg(\frac{\sum_{|e|=1}\omega(X_{k-1},X_{k-1}+e)e^{(\lambda_0+p\bar\lambda)e\cdot e_1}}{\sum_{|e|=1}\omega(X_{k-1},X_{k-1}+e)e^{\lambda_0 e\cdot e_1}}\bigg)+R_{\omega,\lambda_0}(\bar\lambda,t)\\&=\log G_{\omega,\lambda_0}(p\bar\lambda,t)+R_{\omega,\lambda_0}(\bar\lambda,t),\end{aligned}$$

with a remainder term

$$R_{\omega,\lambda_0}(\bar\lambda,t)=\sum_{k=1}^t\log\bigg(\frac{\sum_{|e|=1}\omega(X_{k-1},X_{k-1}+e)e^{(\lambda_0+p\bar\lambda)e\cdot e_1}}{\sum_{|e|=1}\omega(X_{k-1},X_{k-1}+e)e^{\lambda_0 e\cdot e_1}}\bigg)-p\log\bigg(\frac{\sum_{|e|=1}\omega(X_{k-1},X_{k-1}+e)e^{\lambda e\cdot e_1}}{\sum_{|e|=1}\omega(X_{k-1},X_{k-1}+e)e^{\lambda_0 e\cdot e_1}}\bigg).$$

After expanding the exponential and then the logarithm as for (3.5), we get

$$\begin{aligned}R_{\omega,\lambda_0}(\bar\lambda,t)&=\sum_{k=1}^t\bigg(p\bar\lambda d_{\omega,\lambda_0}(X_{k-1})+\frac{1}{2}p^2\bar\lambda^2\big(d^{(2)}_{\omega,\lambda_0}(X_{k-1})-d_{\omega,\lambda_0}(X_{k-1})^2\big)+o\big(p^2\bar\lambda^2\big)\bigg)\\&\quad-p\bigg(\bar\lambda d_{\omega,\lambda_0}(X_{k-1})+\frac{1}{2}\bar\lambda^2\big(d^{(2)}_{\omega,\lambda_0}(X_{k-1})-d_{\omega,\lambda_0}(X_{k-1})^2\big)+o\big(\bar\lambda^2\big)\bigg)\\&\le\big(p^2-p\big)\alpha+o(1)\le p^2\alpha+1\end{aligned}$$

for $|\bar{\lambda}|$ smaller than some $\eta > 0$. For such a choice of λ,

$$\mathbb{E}_{\lambda_0}\big[G_{\omega,\lambda_0}(\bar{\lambda}, t)^p\big] \le \mathbb{E}_{\lambda_0}\big[G_{\omega,\lambda_0}(p\bar{\lambda}, t)\big]e^{p^2\alpha+1} = e^{p^2\alpha+1}. \tag{3.6}$$

Consequently, $(G_{\omega,\lambda_0}(\bar{\lambda}, t))_{|\bar{\lambda}|\le\eta}$ is uniformly bounded in $L^p(\mathbb{P}_{\lambda_0})$. Since this implies convergence of expectations, we get for the density (3.5)

$$\frac{dP_{\omega,\lambda}}{dP_{\omega,\lambda_0}}(X_k; 0 \le k \le t) \xrightarrow[\bar{\lambda}\to 0]{d} \exp\left\{\widehat{M}_\alpha \cdot e_1 - \frac{1}{2}\mathbb{E}_{\lambda_0}\big[(\widehat{M}_\alpha \cdot e_1)^2\big]\right\}$$

under $\mathbb{P}_{\lambda_0}$. By Proposition 3.1, we have also the weak convergence of the product

$$\frac{\bar{\lambda}}{\alpha}\big(X_t - t \cdot v(\lambda_0)\big)\frac{dP_{\omega,\lambda}}{dP_{\omega,\lambda_0}}(X_s; 0 \le s \le t) \xrightarrow[\bar{\lambda}\to 0]{d} \frac{1}{\alpha}\widehat{N}_\alpha \exp\left\{\widehat{M}_\alpha \cdot e_1 - \frac{1}{2}\mathbb{E}_{\lambda_0}\big[(\widehat{M}_\alpha \cdot e_1)^2\big]\right\}.$$

Moreover, this product is by Lemma 3.2 and the calculations above bounded in $L^2(\mathbb{P}_{\lambda_0})$. In particular, it is uniformly integrable and so the expectations converge as well,

$$\mathbb{E}_\lambda\left[\frac{\bar{\lambda}}{\alpha}\big(X_t - t \cdot v(\lambda_0)\big)\right] \xrightarrow[\bar{\lambda}\to 0]{} \frac{1}{\alpha}\mathbb{E}_{\lambda_0}\left[\widehat{N}_\alpha \exp\left\{\widehat{M}_\alpha \cdot e_1 - \frac{1}{2}\mathbb{E}_{\lambda_0}\big[(\widehat{M}_\alpha \cdot e_1)^2\big]\right\}\right].$$

By Girsanov's theorem, the limit is equal to the covariance $\mathrm{Cov}_{\lambda_0}(M, N)e_1$ (recalling $M = \widehat{M}_1$, $N = \widehat{N}_1$). $\square$

3.3. *Proof of Theorem* 3.2

Define $\gamma_n = \mathbb{E}_\lambda[\tau_n]$ and for $t > 0$ fixed, let $n \ge 0$ be such that $\gamma_n \le t < \gamma_{n+1}$. Then

$$\begin{aligned}\left\|\frac{1}{t}\mathbb{E}_\lambda[X_t] - \frac{1}{\gamma_n}\mathbb{E}_\lambda[X_{\gamma_n}]\right\| &\le \frac{1}{t}\big\|\mathbb{E}_\lambda[X_t] - \mathbb{E}_\lambda[X_{\gamma_n}]\big\| + \mathbb{E}_\lambda\big[\|X_{\gamma_n}\|\big]\left|\frac{1}{t} - \frac{1}{\gamma_n}\right| \\ &\le \frac{\gamma_{n+1} - \gamma_n}{\gamma_n} + \gamma_n\frac{t - \gamma_n}{t\gamma_n} \\ &\le 2\frac{\gamma_{n+1} - \gamma_n}{\gamma_n} \le \frac{c}{n},\end{aligned}$$

by the moment bounds of Lemma 3.2. Next, using the 1-dependence of $(\tau_k - \tau_{k-1})_{k\ge1}$ we have,

$$\begin{aligned}\frac{1}{\gamma_n}\big\|\mathbb{E}_\lambda[X_{\gamma_n}] - \mathbb{E}_\lambda[X_{\tau_n}]\big\| &\le \frac{1}{\gamma_n}\mathbb{E}_\lambda\big[(\tau_n - \gamma_n)^2\big]^{1/2} \\ &\le \frac{1}{\gamma_n}\mathbb{E}_\lambda\left[\left(\sum_{k=1}^n(\tau_k - \tau_{k-1}) - (\gamma_k - \gamma_{k-1})\right)^2\right]^{1/2} \\ &\le C\frac{\sqrt{n}}{\gamma_n} \le \frac{C}{\sqrt{n}}.\end{aligned}$$

By the law of large numbers and stationarity of the inter-regeneration times, the speed is given by

$$v(\lambda) = \frac{\mathbb{E}_\lambda[X_{\tau_2} - X_{\tau_1}]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]} = \frac{\mathbb{E}_\lambda[X_{\tau_n} - X_{\tau_1}]}{\mathbb{E}_\lambda[\tau_n - \tau_1]} = \frac{\mathbb{E}_\lambda[X_{\tau_n}] - \mathbb{E}_\lambda[X_{\tau_1}]}{\gamma_n - \gamma_1}$$

such that we have

$$\left\|\frac{\mathbb{E}_\lambda[X_{\tau_n}]}{\gamma_n} - v(\lambda)\right\| \le \big\|\mathbb{E}_\lambda[X_{\tau_n}]\big\|\left|\frac{1}{\gamma_n} - \frac{1}{\gamma_n - \gamma_1}\right| + \frac{1}{\gamma_n - \gamma_1}\big\|\mathbb{E}_\lambda[X_{\tau_1}]\big\| \le \gamma_n\frac{\gamma_1}{\gamma_n(\gamma_n - \gamma_1)} + \frac{\gamma_1}{\gamma_n - \gamma_1} \le \frac{C}{n}.$$

Putting the above estimates together, we get

$$\left\| \frac{1}{t}\mathbb{E}_\lambda[X_t] - v(\lambda) \right\| \le \frac{C}{\sqrt{n}}. \tag{3.7}$$

Recall that we set $\bar{\lambda} = \lambda - \lambda_0$ and $t = \alpha/\bar{\lambda}^2$. Hence $t < \gamma_{n+1} \le cn$, implying

$$\frac{1}{\sqrt{n}} \le \frac{c}{\sqrt{t}} = \frac{c\bar{\lambda}}{\sqrt{\alpha}}.$$

This and the inequality (3.7) implies the estimate of Theorem 3.2. □

4. Monotonicity

4.1. *Proof of Theorem* 1.1

Let us assume already $\delta \le \frac{1}{2}$. By Fact 2, it is possible to choose $\lambda_c < \infty$, such that the speed is monotone on $[\lambda_c, \infty)$, so it suffices to show (strict) monotonicity of v_1 on $[0, \lambda_c]$. We do this by showing that the derivative on this compact interval is strictly positive. More precisely, we compare v_1' with $\bar{v}_1'$, where

$$\bar{v}_1(\lambda) = \frac{e^\lambda - e^{-\lambda}}{e^\lambda + e^{-\lambda} + 2d - 2}$$

is the speed of the random walk in a homogeneous environment $\bar{\omega}$, where all conductances equal 1. Since $\bar{v}_1'$ is greater than some positive ε_0 on $[0, \lambda_c]$, positivity of v_1' follows from

$$\sup_{\lambda \in [0,\lambda_c]} \left| v_1'(\lambda) - \bar{v}_1'(\lambda) \right| < \varepsilon_0 \tag{4.1}$$

for δ close enough to 0. In Section 2 we constructed a coupling $(X_n^{(\omega,\lambda)}, X_n^{(\bar{\omega},\lambda)})_n$ between the random walk in an original environment ω and a random walk in the homogeneous environment $\bar{\omega}$. To keep the notation simpler, we denote $X_n^{(\omega,\lambda)}$ again by X_n and $X_n^{(\bar{\omega},\lambda)}$ by $\bar{X}_n$. Furthermore, define analogously to (1.5) and (1.6) the processes $\bar{M}_n$ and $\bar{N}_n$ in the homogeneous environment. (Of course, $\bar{M}_n = \bar{N}_n$.) The coupling guarantees then

$$P(X_n - X_{n-1} \ne \bar{X}_n - \bar{X}_{n-1}) \le C\delta, \tag{4.2}$$

so if δ is sufficiently small, the two processes will take the same steps most of the time. By Theorem 1.3 and the moment bounds in Lemma 3.2, we have

$$\begin{aligned} &v_1'(\lambda) - \bar{v}_1'(\lambda) \\ &\quad = \lim_{n\to\infty} \frac{1}{n}\big[\mathrm{Cov}_\lambda(M_n, N_n)_{1,1} - \mathrm{Cov}_\lambda(\bar{M}_n, \bar{N}_n)_{1,1}\big] \\ &\quad = \lim_{n\to\infty} \frac{1}{n}\big[\mathrm{Cov}_\lambda(M_n - \bar{M}_n, N_n - \bar{N}_n)_{1,1} + \mathrm{Cov}_\lambda(M_n - \bar{M}_n, \bar{N}_n)_{1,1} + \mathrm{Cov}_\lambda(\bar{M}_n, N_n - \bar{N}_n)_{1,1}\big]. \end{aligned}$$

By the Cauchy–Schwarz inequality,

$$\begin{aligned} \left|v_1'(\lambda) - \bar{v}_1'(\lambda)\right| &\le \limsup_{n\to\infty} \frac{1}{n}\big[\mathrm{Var}_\lambda(M_n - \bar{M}_n)_{1,1}^{1/2}\, \mathrm{Var}_\lambda(\bar{N}_n)_{1,1}^{1/2}\big] \\ &\quad + \limsup_{n\to\infty} \frac{1}{n}\big[\mathrm{Var}_\lambda(N_n - \bar{N}_n)_{1,1}^{1/2}\big(\mathrm{Var}_\lambda(M_n - \bar{M}_n)_{1,1}^{1/2} + \mathrm{Var}_\lambda(\bar{M}_n)_{1,1}^{1/2}\big)\big]. \end{aligned} \tag{4.3}$$

We will first show the following bounds (which are in fact actual limits):

$$\limsup_{n\to\infty} \frac{1}{n} \operatorname{Var}_\lambda(\bar{N}_n)_{1,1} \le C \tag{4.4}$$

$$\limsup_{n\to\infty} \frac{1}{n} \operatorname{Var}_\lambda(\bar{M}_n)_{1,1} \le C \tag{4.5}$$

$$\limsup_{n\to\infty} \frac{1}{n} \operatorname{Var}_\lambda(M_n - \bar{M}_n)_{1,1} \le C\delta. \tag{4.6}$$

The first two bounds (4.4) and (4.5) follow since $\bar{N}_n = \bar{M}_n$ is a process in the homogeneous environment with iid increments uniformly bounded in λ and δ.

For (4.6), observe that $(M_n - \bar{M}_n)\cdot e_1$ is again a martingale with

$$\begin{aligned}
&\mathbb{E}_\lambda\bigl[\bigl((M_n - \bar{M}_n)\cdot e_1 - (M_{n-1} - \bar{M}_{n-1})\cdot e_1\bigr)^2\bigr] \\
&\quad \le 2\mathbb{E}_\lambda\bigl[\bigl((X_n - X_{n-1})\cdot e_1 - (\bar{X}_n - \bar{X}_{n-1})\cdot e_1\bigr)^2\bigr] + 2\mathbb{E}_\lambda\bigl[\bigl(d_{\omega,\lambda}(X_{n-1})\cdot e_1 - d_{\bar{\omega},\lambda}(\bar{X}_{n-1})\cdot e_1\bigr)^2\bigr],
\end{aligned}$$

where $d_{\omega,\lambda}(x) = E^x_{\omega,\lambda}[X_1 - X_0]$. By (4.2), the first term is of order δ. We have

$$\bigl\|d_{\omega,\lambda}(x) - d_{\bar{\omega},\lambda}(x)\bigr\| \le C\delta, \tag{4.7}$$

so that the second term is of order at most δ as well. Consequently,

$$\limsup_{n\to\infty} \frac{1}{n}\mathbb{E}_\lambda\bigl[\bigl((M_n - \bar{M}_n)\cdot e_1\bigr)^2\bigr] \le C\delta.$$

With the bounds (4.4) to (4.6), we may bound (4.3) as

$$\bigl|v_1'(\lambda) - \bar{v}_1'(\lambda)\bigr| \le C\sqrt{\delta} + C\limsup_{n\to\infty} \frac{1}{n} \operatorname{Var}_\lambda(N_n - \bar{N}_n)^{1/2}_{1,1}. \tag{4.8}$$

To bound the remaining variance, we decompose

$$(N_n - \bar{N}_n)\cdot e_1 = \bigl(X_n - \bar{X}_n - n\bigl(v(\lambda) - \bar{v}(\lambda)\bigr)\bigr)\cdot e_1 = (M_n - \bar{M}_n)\cdot e_1 + Z_n, \tag{4.9}$$

where

$$Z_n = \sum_{k=0}^{n-1}\bigl(d_{\omega,\lambda}(X_k) - d_{\bar{\omega},\lambda}(X_k)\bigr)\cdot e_1 - \bigl(v_1(\lambda) - \bar{v}_1(\lambda)\bigr). \tag{4.10}$$

Note that we obtain from (4.9) and (3.3) that

$$\mathbb{E}_\lambda[Z_{\tau_2} - Z_{\tau_1}] = 0. \tag{4.11}$$

We already know from (4.6) that the difference of the martingales in (4.9) is nicely bounded, so we just need to bound (the $n\to\infty$ limit of) $\mathbb{E}_\lambda[(Z_n)^2]/n$. Below we estimate this limit.

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\mathbb{E}_\lambda\bigl[(Z_n)^2\bigr] &= \lim_{n\to\infty} \frac{\mathbb{E}_\lambda[(Z_{\tau_n})^2]}{\mathbb{E}_\lambda[\tau_n]} \\
&= \frac{\mathbb{E}_\lambda[(Z_{\tau_2} - Z_{\tau_1})^2] + 2\mathbb{E}_\lambda[(Z_{\tau_3} - Z_{\tau_2})(Z_{\tau_2} - Z_{\tau_1})]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]} \\
&\le 3\frac{\mathbb{E}_\lambda[(Z_{\tau_2} - Z_{\tau_1})^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]}.
\end{aligned} \tag{4.12}$$

For the first equality, we just notice that the limit does not change when taken along the subsequence of regeneration times. For the second equality we use the fact that $(Z_{\tau_n} - Z_{\tau_{n-1}}, \tau_n - \tau_{n-1})$ is a stationary 1-dependent sequence, and for the inequality we just use Cauchy–Schwartz. Furthermore, let

$$\xi_n = \sum_{k=0}^{n-1} \bigl(d_{\omega,\lambda}(X_k) - d_{\bar{\omega},\lambda}(X_k)\bigr) \cdot e_1. \tag{4.13}$$

Then (4.11) implies that the speed may be written as

$$v_1(\lambda) - \bar{v}_1(\lambda) = \frac{\mathbb{E}_\lambda[\xi_{\tau_2} - \xi_{\tau_1}]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]}. \tag{4.14}$$

For the last line in (4.12), we obtain then by Jensen's inequality

$$\begin{aligned}
\frac{\mathbb{E}_\lambda[(Z_{\tau_2} - Z_{\tau_1})^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]} &\le 2\frac{\mathbb{E}_\lambda[(\xi_{\tau_2} - \xi_{\tau_1})^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]} + 2\frac{\mathbb{E}_\lambda[(\tau_2 - \tau_1)^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]}\bigl(v_1(\lambda) - \bar{v}_1(\lambda)\bigr)^2 \\
&= 2\frac{\mathbb{E}_\lambda[(\xi_{\tau_2} - \xi_{\tau_1})^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]} + 2\frac{\mathbb{E}_\lambda[(\tau_2 - \tau_1)^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]}\left(\frac{\mathbb{E}_\lambda[(\xi_{\tau_2} - \xi_{\tau_1})]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]}\right)^2 \\
&\le 2\frac{\mathbb{E}_\lambda[(\xi_{\tau_2} - \xi_{\tau_1})^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]}\left(1 + \frac{\mathbb{E}_\lambda[(\tau_2 - \tau_1)^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]^2}\right).
\end{aligned} \tag{4.15}$$

From the moment bounds in Lemma 3.1 and (3.2), we conclude the uniform bound

$$\frac{\mathbb{E}_\lambda[(Z_{\tau_2} - Z_{\tau_1})^2]}{\mathbb{E}_\lambda[\tau_2 - \tau_1]} \le C\lambda^2 \mathbb{E}_\lambda\bigl[(\xi_{\tau_2} - \xi_{\tau_1})^2\bigr]. \tag{4.16}$$

With the estimates (4.12) and (4.16), the inequality (4.8) implies

$$\bigl|v_1'(\lambda) - \bar{v}_1'(\lambda)\bigr| \le C\sqrt{\delta} + C\mathbb{E}_\lambda\bigl[\lambda^2(\xi_{\tau_2} - \xi_{\tau_1})^2\bigr]^{1/2}. \tag{4.17}$$

The uniform bound (4.7) gives

$$\mathbb{E}_\lambda\bigl[\lambda^2(\xi_{\tau_2} - \xi_{\tau_1})^2\bigr] \le C\delta^2 \mathbb{E}_\lambda\bigl[\lambda^2(\tau_2 - \tau_1)^2\bigr] \le C\frac{\delta^2}{\lambda^2}, \tag{4.18}$$

where we used (3.1) for the last inequality. Of course, this bound blows up near $\lambda = 0$, but it shows that for any $\lambda_0 > 0$, the right hand side of (4.17) may be bounded uniformly for $\lambda \in [\lambda_0, \lambda_c]$ and then

$$\lim_{\delta \to 0} \sup_{\lambda \in [\lambda_0, \lambda_c]} \bigl|v_1'(\lambda) - \bar{v}_1'(\lambda)\bigr| = 0.$$

So for any $\lambda_0 > 0$, we can choose $\delta = \delta(\lambda_0)$ small enough, so that the speed is monotone on $[\lambda_0, \infty)$. It remains to show that for some ellipticity constant $\delta > 0$, we have monotonicity on the whole range of $[0, \infty)$.

Now suppose there are environment measures compatible with our a priori bound $\delta \le \frac{1}{2}$ such that the speed is not monotone on $[0, \infty)$. If there is some $\delta' > 0$ such that none of these measures satisfies the uniform ellipticity assumption with δ', we may just choose δ_0 in Theorem 1.1 accordingly to exclude these measures. Otherwise, there exists a sequence $P^{(n)}$ of environment measures with ellipticity constants $\delta_n \to 0$ and such that the speed is not monotone. In this case we may find a sequence of $\lambda_n > 0$ with $|v_{n,1}'(\lambda_n) - \bar{v}_{n,1}'(\lambda_n)| \ge \varepsilon_0$, where $v_{n,1}$ is the first coordinate of the speed under $P^{(n)}$. By the bounds (4.17) and (4.18), we have necessarily $\lambda_n \to 0$ and the left hand side of (4.18) does not vanish. To complete the proof it remains to show that such a sequence cannot exist. The following lemma, combined with (4.17), shows indeed that the existence of such a sequence is impossible.

Lemma 4.1. *For any sequence of environment measures $P^{(n)}$ with ellipticity constants $\delta^{(n)} \to 0$ and any sequence λ_n with $\lambda_n \to 0$,*

$$\lim_{n\to\infty} \mathbb{E}_{\lambda_n}^{(n)}\big[\lambda_n^2(\xi_{\tau_2} - \xi_{\tau_1})^2\big] = 0.$$

Proof. To simplify notation, let us drop some of the indices n, in particular we write λ for λ_n. We have for $i = 1, 2$

$$\begin{aligned}\mathbb{E}_\lambda^{(n)}\big[\lambda^2(\xi_{\tau_i})^2\big] &\le \sum_{N=1}^{\infty} \mathbb{E}_\lambda^{(n)}\big[\lambda^2\big(\xi_{N/\lambda^2}^*\big)^2 \mathbb{1}_{\{(N-1)/\lambda^2 \le \tau_i < N/\lambda^2\}}\big] \\ &\le \sum_{N=1}^{\infty} \mathbb{E}_\lambda^{(n)}\big[\lambda^3|\xi_{N/\lambda^2}^*|^3\big]^{2/3} \mathbb{P}_\lambda^{(n)}\big(\tau_i \ge N/\lambda^2\big)^{1/3},\end{aligned}$$

with

$$\xi_{N/\lambda^2}^* = \max_{0\le k\le N/\lambda^2} \xi_k.$$

By the moment bound for τ_i,

$$\mathbb{E}_\lambda^{(n)}\big[\lambda^2(\xi_{\tau_i})^2\big] \le C \sum_{N=1}^{\infty} \mathbb{E}_\lambda^{(n)}\big[\lambda^3\big|\xi_{N/\lambda^2}^*\big|^3\big]^{2/3} e^{-cN}.$$

Using the decomposition of $\xi_k = (M_k - \bar{M}_k) \cdot e_1 + (X_k - \bar{X}_k) \cdot e_1$ into a martingale term with bounded increments and the process X_k, Doob's inequality and the bound in Lemma 3.2 implies

$$\mathbb{E}_\lambda^{(n)}\big[\lambda^4\big|\xi_{N/\lambda^2}^*\big|^4\big] \le CN^4, \tag{4.19}$$

such that by the dominated convergence theorem the assertion of the lemma will follow once we show that for every N,

$$\lim_{n\to\infty} \mathbb{E}_\lambda^{(n)}\big[\lambda^3\big|\xi_{N/\lambda^2}^*\big|^3\big] = 0.$$

We write the expectation with respect to the unbiased measure,

$$\mathbb{E}_\lambda^{(n)}\big[\lambda^3\big|\xi_{N/\lambda^2}^*\big|^3\big] = \mathbb{E}_0^{(n)}\big[\lambda^3\big|\xi_{N/\lambda^2}^*\big|^3 G\big(\omega^{(n)}, \lambda, N/\lambda^2\big)\big],$$

with

$$G(\omega, \lambda, m) = \frac{dP_{\omega,\lambda}}{dP_{\omega,0}}(X_k; 0 \le k \le m),$$

and $\omega^{(n)}$ distributed according to $P^{(n)}$. We know that

$$G\big(\omega^{(n)}, \lambda, N/\lambda^2\big) = \exp\bigg(\lambda M_{N/\lambda^2} \cdot e_1 - \frac{1}{2}\mathbb{E}_0^{(n)}\big[(\lambda M_{N/\lambda^2} \cdot e_1)^2\big] + o(\lambda)\bigg),$$

with an error term uniform in δ. Since δ and the distribution of $\omega^{(n)}$ is now varying with λ, M_{N/λ^2} is now a triangular array of martingales. Thanks to the fact that all increments are uniformly (in δ and λ) bounded, the CLT for arrays of martingales yields

$$G\big(\omega^{(n)}, \lambda, N/\lambda^2\big) \xrightarrow[n\to\infty]{d} e^{\widehat{M}_N \cdot e_1 - \frac{1}{2}E[(\widehat{M}_N \cdot e_1)^2]},$$

with $\widehat{M}_N$ a Gaussian random variable. Again, this convergence is complemented by a good moment bound, see (3.6),

$$\mathbb{E}_0^{(n)}\big[G\big(\omega^{(n)},\lambda,N/\lambda^2\big)^p\big]\le e^{p^2\frac{N}{2}+1}$$

for all n and $p\ge 1$. Therefore, it suffices to show

$$\lambda\xi^*_{N/\lambda^2}\xrightarrow[n\to\infty]{}0 \tag{4.20}$$

in probability. Until now we tacitly ignored that $d_{\omega,\lambda}(x)$ in the definition of $\xi_n=\xi_n(\lambda)$ depends on λ, but by the bound

$$\big\|\big(d_{\omega,\lambda}(x)-d_{\bar{\omega},\lambda}(x)\big)-\big(d_{\omega,0}(x)-d_{\bar{\omega},0}(x)\big)\big\|\le C\delta\lambda$$

we have

$$\lambda\big|\xi^*_{N/\lambda^2}(\lambda)-\xi^*_{N/\lambda^2}(0)\big|\le CN\delta.$$

Therefore, it suffices to show that $\lambda\xi^*_{N/\lambda^2}(0)$ goes in probability to zero as n goes to infinity. Recall that since for $\lambda=0$ the local drift in the environment $\bar{\omega}$ is zero, i.e. $d_{\bar{\omega},0}(x)=0,\forall x$, we get in fact

$$\xi_n(0)=\sum_{k=0}^{n-1}d_{\omega,0}(X_k)\cdot e_1.$$

Lemma 4.2 (with L of that lemma set to be N/λ^2) below shows that

$$\mathbb{E}_0^{(n)}\big[\xi^*_{N/\lambda^2}(0)^2\big]\le CN\lambda^{-2}\delta,$$

so

$$\mathbb{E}_0^{(n)}\big[\big(\lambda\xi^*_{N/\lambda^2}(0)\big)^2\big]\le CN\delta$$

which goes to zero as n goes to infinity since then $\delta=\delta^{(n)}\to 0$. □

The next lemma is now all that is missing. The main idea of its proof originated, to the best of our knowledge, in [7], and was further developed by [13] (Proposition 3.3) and [14] (Lemma 2.4).

Lemma 4.2. *There exists a constant $C>0$ depending only on the dimension, such that for all $L\ge 1$ and $\delta\le\frac{1}{2}$, we have, with $d_\omega(\cdot)=d_{\omega,0}(\cdot)$,*

$$\mathbb{E}_0\Bigg[\sup_{0\le n\le L}\Bigg\|\sum_{k=0}^{n-1}d_\omega(X_k)\Bigg\|^2\Bigg]\le CL\delta.$$

Proof. Recall that the environment measure Q with

$$\frac{dQ}{dP}(\omega)=Z^{-1}\sum_{|e|=1}\omega(0,e)$$

is stationary, reversible and ergodic for the process $(\widehat{\omega}_n)_n$ of the environment seen from the particle (see [14] and [12] for the definition of $(\widehat{\omega}_n)_n$ and some properties). If $\delta\le\frac{1}{2}$, the density satisfies $c\le|\frac{dQ}{dP}(\omega)|\le C$ with positive constants c, C depending only on the dimension. Therefore we may consider expectation with respect to $Q\times P_\omega$, which we denote by $\mathbb{E}_Q$. Under this measure,

$$M_n=(X_n-\bar{X}_n)-(X_0-\bar{X}_0)-\sum_{k=0}^{n-1}d_\omega(X_k)$$

is a martingale with respect to the filtration $\mathcal{F}_n = \sigma(\{\widehat{\omega}_0, \dots, \widehat{\omega}_n\})$. Since by time reversal, for any $n \geq 1$, the sequence

$$\big((X_1 - X_0) - (\bar{X}_1 - \bar{X}_0), \dots, (X_n - X_{n-1}) - (\bar{X}_n - \bar{X}_{n-1}), \widehat{\omega}_0, \dots, \widehat{\omega}_n\big)$$

has the same distribution as

$$\big((\bar{X}_n - \bar{X}_{n-1}) - (X_n - X_{n-1}), \dots, (\bar{X}_1 - \bar{X}_0) - (X_1 - X_0), \widehat{\omega}_n, \dots, \widehat{\omega}_0\big)$$

under $Q_0 \times P_\omega$, we have that

$$M_n^- = (X_{L-n} - \bar{X}_{L-n}) - (X_L - \bar{X}_L) - \sum_{k=0}^{n-1} d_\omega(X_{L-k})$$

is a martingale with respect to the filtration $\mathcal{F}_n^- = \sigma(\{\widehat{\omega}_L, \dots, \widehat{\omega}_{L-n}\})$. Noting that

$$M_L^- - M_{L-n}^- = (X_0 - \bar{X}_0) - (X_n - \bar{X}_n) - \sum_{k=1}^{n} d_\omega(X_k),$$

we get

$$M_n + M_L^- - M_{L-n}^- = -2\sum_{k=0}^{n-1} d_\omega(X_k) + d_\omega(X_0) - d_\omega(X_n).$$

Therefore,

$$\mathbb{E}_Q\left[\sup_{0\leq n\leq L}\left\|\sum_{k=0}^{n-1} d_\omega(X_k)\right\|^2\right] \leq \frac{1}{4}\mathbb{E}_Q\Big[\sup_{0\leq n\leq L}\big\|M_n + M_L^- - M_{L-n}^- - d_\omega(X_0) + d_\omega(X_n)\big\|^2\Big].$$

The lemma follows then from Doob's inequality, since $|d_\omega(x)| \leq C\delta$ and

$$\mathbb{E}_Q\big[\|M_L\|^2\big] + \mathbb{E}_Q\big[\big\|M_L^-\big\|^2\big] \leq CL\delta. \qquad \square$$

4.2. *Proof of Theorem* 1.2

The proof follows the arguments of [5], where the speed of biased random walk on a percolation cluster is studied. Note that the environment measure with

$$P\big(\omega(0,e) = 1\big) = p = 1 - P\big(\omega(0,e) = \kappa\big)$$

generates a percolation graph consisting of the edges with conductance 1, connected by κ-edges. So if $p > \frac{1}{2}$ and κ small enough, we would expect the random walk to behave like the random walk on the percolation cluster for most times, with short excursions along κ-edges. In analogy with the percolation case, we say in this section that an edge $\{x, y\}$ is open if $\omega(x, y) = 1$ and (infinite) cluster will mean the (infinite) cluster connected by open edges.

We choose a bias λ_1, such that the random walk on the percolation cluster has a positive speed and show

$$v_1(\lambda_1) \geq c_0 \tag{4.21}$$

for a positive c_0 independent of κ. On the other hand, for a larger bias λ_2, chosen such that the random walk on the percolation cluster has zero speed, we show

$$v_1(\lambda_2) \leq c_0/2 \tag{4.22}$$

for κ sufficiently small. The combination of these two bounds yields the statement of Theorem 1.2.

4.2.1. *A lower bound for* $v_1(\lambda_1)$

Denote the infinite cluster connected by open edges by I.

Definition 4.1. A point $x \in \mathbb{Z}^2$ is good, if there exists an infinite path $x = x_0, x_1, x_2, \dots$ such that for all $k \geq 1$

(i) $|(x_k - x_{k-1}) \cdot e_2| = 1$ and $(x_k - x_{k-1}) \cdot e_1 = 1$,
(ii) the edges $\{x_{k-1}, x_{k-1} + e_1\}, \{x_{k-1} + e_1, x_k\}$ are open.

Let J be the set of good vertices. We say a vertex x is bad, if $x \in I$ and x is not good. Connected components of $I \setminus J$ are called traps. For a vertex x, let $T(x)$ be the trap containing x (being empty if x is good). The length of the trap of x is

$$L(x) = \sup\{(y - z) \cdot e_1 : y, z \in T(x)\}$$

and the width is

$$W(x) = \sup\{(y - z) \cdot e_2 : y, z \in T(x)\}.$$

If $T(x)$ is empty, then we take $L(x) = W(x) = 0$. The following estimate is Lemma 1 in [5].

Lemma 4.3. *For every* $p \in (\frac{1}{2}, 1)$ *there exists* $\alpha = \alpha(p)$ *such that* $P(L(0) \geq n) \leq \alpha^n$ *and* $P(W(0) \geq n) \leq \alpha^n$ *for every* n. *Further*, $\lim_{p \to 1} \alpha(p) = 0$.

Let $\mathcal{H}(n)$ be the σ-algebra generated by the history of the random walk until time n, i.e., $\mathcal{H}(n) = \sigma(\{X_0 = 0, X_1, X_2, \dots, X_n\})$. Let $P_{\omega,\lambda}^{\mathcal{H}(n)}$ be the conditional distribution of $P_{\omega,\lambda}$ given $\mathcal{H}(n)$, and $\mathbb{P}_{\lambda}^{\mathcal{H}(n)}$ be the conditional distribution of $\mathbb{P}_{\lambda}$ given $\mathcal{H}(n)$. Define $\tau_n(h) = \min\{i > n : X_i \cdot e_1 = h\}$. The following estimate is essential in the proof of the lower bound.

Lemma 4.4. *There exists* $D' = D'(\lambda)$ *such that for every* $\ell \geq 1$ *and for every configuration* ω *such that* x *is a good point*,

$$P_{\omega,\lambda}^{\mathcal{H}(n)}\bigl(\tau_n(x \cdot e_1 - \ell) \leq \tau_n(x \cdot e_1 + \ell/3) | X_n = x\bigr) < D' e^{-\lambda\ell/3}.$$

Proof. Consider the box $B = x + [-\ell, \ell/3] \times [-e^{\lambda\ell}, e^{\lambda\ell}]$ with right face $B^+ = x + \{\ell/3\} \times [-e^{\lambda\ell}, e^{\lambda\ell}]$. From the general theory of electrical networks, see [9] or [17], we have the inequality

$$P_{\omega,\lambda}^{\mathcal{H}(n)}\bigl(\tau_n(x \cdot e_1 - \ell) \leq \tau_n(x \cdot e_1 + \ell/3) | X_n = x\bigr) \leq \frac{C_{x,\partial B \setminus B^+}}{C_{x,B^+}},$$

where $C_{x,A}$ denoted the effective conductance between a point x and a set A (see also Fact 2 in [5]). The conductance C_{x,B^+} is bounded from below by the conductance of a good path from x to B^+, which is at least $D_1 e^{\lambda 2x \cdot e_1}$ for some $D_1 = D_1(\lambda)$. Furthermore, we have the upper bound

$$C_{x,\partial B \setminus B^+} \leq C_{x,\partial B^-} + C_{x,\partial B_1} + C_{x,\partial B_2},$$

where

$$B^- = x + \{-\ell\} \times \bigl[-e^{\lambda\ell}, e^{\lambda\ell}\bigr],$$

$$B_1 = x + [-\ell, \ell/3] \times \bigl\{-e^{\lambda\ell}\bigr\},$$

$$B_2 = x + [-\ell, \ell/3] \times \bigl\{e^{\lambda\ell}\bigr\}.$$

The effective conductance $C_{x,\partial B^-}$ is bounded from above by the sum of the edge weights between z and $z+e_1$, for $z \in B^-$. But for every such z, the weight is

$$\omega(z, z+e_1)e^{\lambda(2z\cdot e_1+1)} \le e^{\lambda(2x\cdot e_1-2\ell+1)}.$$

There are at most $2e^{\lambda\ell}+1$ such edges. Therefore $C_{x,\partial B^-} \le D_2 e^{\lambda(2x\cdot e_1-\ell)}$ for some $D_2 = D_2(\lambda)$. Finally, the Nash–Williams inequality gives

$$C_{x,B_i} \le e^{-\lambda\ell} \sum_{i=x\cdot e_1-\ell}^{x\cdot e_1+\ell/3} e^{2\lambda(i+1)} \le D_3 e^{\lambda(2x\cdot e_1-\ell/3)}$$

for some $D_3 = D_3(\lambda)$. Combining the bounds for the effective conductances, we get the desired bound for the exit probability. □

Let $G(x)$ be the event that x is a good point. We call a time point n a fresh epoch, if $(X_n - X_k)\cdot e_1 > 0$ for all $k < n$ and let $F(n)$ be the event that n is a fresh epoch. From the bound in Lemma 4.4, we get the following inequalities (Lemma 3 and Lemma 4 in [5]). In the following, take p so close to 1 that $\alpha(p)$ in Lemma 4.3 is less than 1. Then there exists a constant $D = D(\lambda, p)$ such that

$$\mathbb{P}_\lambda^{\mathcal{H}(n)}\big(\text{ there is an } m \ge n \text{ such that } (X_m - X_n)\cdot e_1 \le -\ell | F(n), G(X_n)\big) \le De^{-\lambda\sqrt{\ell}/D}, \quad \mathbb{P}_\lambda\text{-a.s.} \tag{4.23}$$

Let $\tau_n'(h)$ be the first fresh epoch later than n, such that the random walk hits a good point whose first coordinate is larger or equal to h. Then, there exists a constant $K = K(\lambda, p)$ such that for any $\ell \ge 1$

$$\mathbb{P}_\lambda^{\mathcal{H}(n)}\Big(\tau(X_n\cdot e_1 - \ell) < \tau_n'(X_n\cdot e_1 + \ell/6)\Big|G(X_n), \max_{0\le i\le n}(X_i - X_n)\cdot e_1 < \sqrt{\ell}\Big) \le Ke^{-\lambda\sqrt{\ell}/K}, \tag{4.24}$$

$\mathbb{P}_\lambda$-almost surely. In particular,

$$\mathbb{P}_\lambda^{\mathcal{H}(n)}\big(\tau(X_n\cdot e_1 - \ell) < \tau_n'(X_n\cdot e_1 + \ell/6) | F(n), G(X_n)\big) \le Ke^{-\lambda\sqrt{\ell}/K}, \tag{4.25}$$

$\mathbb{P}_\lambda$-almost surely. From these bounds, the following lower bound for the speed is proven. Note that the constant is independent of κ.

Lemma 4.5. *For λ sufficiently small, there exists a constant $C = C(p)$ such that*

$$\mathbb{P}_\lambda\big(X_n\cdot e_1 < Cn^{1/10}\big) \le Cn^{-2}.$$

Let us highlight the only change necessary in the proof given in [5]: Therein, the Carne-Varopoulos bound

$$P_{\omega,\lambda}^x(X_n = y) \le 2\sqrt{\frac{\pi(y)}{\pi(x)}}\exp\left(-\frac{d(x,y)^2}{2n}\right) \tag{4.26}$$

is applied, with π the reversible measure and $d(\cdot,\cdot)$ the graph distance. On the percolation cluster, it is easy to get a further upper bound, since in this case,

$$e^{\lambda(2x\cdot e_1-1)} \le \pi(x) \le 4e^{\lambda(2x\cdot e_1+1)},$$

as every point x in the cluster is the endpoint of an edge with conductance 1. Of course, the upper bound is still valid in our case, but the lower bound depends on κ if x is surrounded by only κ-edges. To get a lower bound independent of κ, let $J(x)$ be the connected component of points surrounded by κ-edges. If $J(x)$ is empty, we can proceed as in the percolation case. Otherwise, let

$$T_x = \inf\{n \ge 0 : X_n \notin J(x)\}$$

and define for positive integers d_n the events

$$A_n = \{\mathrm{diam}(J(z)) \le d_n \text{ for all } z - x \in [-n, n]^2\},$$

then by Lemma 4.3,

$$P(A_n^c) \le n^2 P(\mathrm{diam}(T(0)) > d_n) \le 2n^2 \alpha^{d_n/2}. \tag{4.27}$$

For an environment $\omega \in A_n$ we have then for the hitting probability

$$P^x_{\omega,\lambda}(X_n = y) \le \sum_{z\in\mathbb{Z}^2} \sum_{1\le m\le n} P^z_{\omega,\lambda}(X_{n-m} = y) P^z_{\omega,\lambda}(T_x = m, X_{T_x} = z). \tag{4.28}$$

On A_n, there are at most d_n^2 points z such that the second probability in the sum is nonzero, and for each such z we have by the Carne-Varopoulos bound

$$P^z_{\omega,\lambda}(X_{n-m} = y) \le 4e^{\lambda((y-z)\cdot e_1+1)} \exp\left(-\frac{d(z,y)^2}{2(n-m)}\right) \le 4e^{\lambda((x-y)\cdot e_1+d_n+1)} \exp\left(-\frac{(d(x,y)-d_n)^2}{2n}\right).$$

Let $d_n = \gamma \log(n)$ for $\gamma = -8/\log(\alpha)$, then for all but finitely many n, A_n occurs. For all $\omega \in A_n$ and $1 \le i < j \le n$ we conclude by the union bound

$$\begin{aligned} P^x_{\omega,\lambda}(X_i \cdot e_1 = X_j \cdot e_1 \text{ but } \|X_i - X_j\| \ge n^{6/10}) &\le 4n^4 d_n^2 e^{\lambda(d_n+1)} \exp\left(-\frac{(n^{6/10}-d_n)^2}{2n}\right) \\ &\le \exp\left(-\frac{1}{5} n^{1/10}\right) \end{aligned}$$

for n sufficiently large, which yields the necessary estimate in [5].

Lemma 4.6. *There exists a constant $c = c(\lambda, p) > 0$ such that*

$$\mathbb{P}_\lambda(X_n \cdot e_1 \ge 1 \textit{ for all } n \ge 1) > c.$$

Proof. Let $\ell_0 = N$ be a positive integer and $\ell_{i+1} = 13\ell_i/12$ for $i \ge 1$. Define recursively the times $t_0 = N$, $t_{i+1} = \tau'_{t_i}(X_{t_i} \cdot e_1 + \ell_i/6)$ and the events

$$A_0 = \{X_N = (N, 0) \text{ and } (N, 0) \text{ is a good point}\}$$

and

$$A_i = \{\tau'_{t_i}(X_{t_i} \cdot e_1 + \ell_i/6) < \tau_{t_i}(X_{t_i} \cdot e_1 - \ell_i)\}.$$

Then $\mathbb{P}_\lambda(A_0) = c_N > 0$ and by (4.25),

$$\mathbb{P}_\lambda(A_i^c) \le K e^{-\lambda\sqrt{\ell_i}/K}.$$

Therefore,

$$\mathbb{P}_\lambda\left(\bigcap_{i=0}^{\infty} A_i\right) \ge c_N(1 - Ce^{-\lambda\sqrt{N}/K}),$$

which is positive for N large enough. When all of the events A_i occur, then $t_i < \infty$ for all i and if $m \ge t_i$,

$$X_m \cdot e_1 > X_{t_i} \cdot e_1 - \ell_i \ge X_{t_0} \cdot e_1 - \ell_0 + \frac{1}{12}\sum_{j=1}^{i-1} \ell_j \ge \frac{N}{12}\left(\frac{13}{12}\right)^{i-1},$$

which implies in particular $X_n \cdot e_1 \ge 1$ for all $n \ge 1$. □

We now introduce a regeneration structure, slightly different from the one used to prove Theorem 1.1. Recall that n is a fresh epoch, if $X_n \cdot e_1 > X_k \cdot e_1$ for all $k < n$. If n is a fresh epoch and additionally, $X_n \cdot e_1 < X_k \cdot e_1$ for all $k > n$, we call n a regeneration and we denote by R_n the nth regeneration time.

For $z \in \mathbb{Z}^2$, let $\omega_z^+ = \{\omega_z(x, y) : x \sim y, x \cdot e_1 \geq z \cdot e_1\}$ be the environment to the right of z. The following lemma is standard in the theory of random walks in random environments, see [20].

Lemma 4.7. *The sequence*

$$\big((X_{R_n+k} - X_{R_n})_{k\geq 0}, \omega_{R_n}^+\big)_{n\geq 1}$$

is stationary and ergodic. Moreover, the distribution of $((X_{R_n+k} - X_{R_n})_{k\geq 0}, \omega_{R_n}^+)$ *is given by the distribution of* $((X_k)_{k\geq 0}, \omega_0^+)$ *under* $\mathbb{P}_\lambda$, *conditioned on* $\{X_n \cdot e_1 \geq 1$ *for all* $n \geq 1\}$.

It follows from Lemma 4.7 that $v(\lambda)$ exists and is nonzero if and only if $\mathbb{E}_\lambda[R_2 - R_1] < \infty$ and in this case

$$v(\lambda) = \frac{\mathbb{E}_\lambda[X_{R_2} - X_{R_1}]}{\mathbb{E}_\lambda[R_2 - R_1]}. \tag{4.29}$$

Since $(X_{R_2} - X_{R_1}) \cdot e_1 \geq 1$, the inequality (4.21) follows then from

$$\mathbb{E}_\lambda[R_2 - R_1] \leq C,$$

with a constant $C = C(\lambda, p)$ independent of κ. This inequality follows by the same arguments as Lemma 8 in [5], making use of Lemma 4.4, Lemma 4.5 and Lemma 4.6.

4.2.2. *An upper bound for* $v_1(\lambda_2)$

The upper bound (4.22) follows from the fact that for small values of κ, the random walk will spend a long time in dead ends of the percolation cluster. To be more precise, let $I(x)$ be the connected component of x connected by open edges (i.e., with conductance 1). We call $x \in \mathbb{Z}^2$ the beginning of a dead end, if x belongs to the infinite cluster to its left, but not to the infinite cluster to its right, i.e., $I(x) \cap \{z : (z - x) \cdot e_1 < 0\}$ is infinite but $I(x) \cap \{z : (z - x) \cdot e_1 \geq 0\}$ is finite. The dead end starting at x is the finite set $I(x) \cap \{z : (z - x) \cdot e_1 \geq 0\}$. Let A be a dead end starting at the origin and $d(A) = \max\{z \cdot e_1 : z \in A\}$ the depth of A. The time spent in A will be denoted by

$$T_A = \inf\{n \geq 1 : X_n \cdot e_1 \leq 0\}. \tag{4.30}$$

If there is no dead end at the origin, set $A = \varnothing$, $d(A) = 0$ and $T_A = 0$. For an environment ω with $\omega(x, y) \in \{\kappa, 1\}$ for $x \sim y$, let $\bar{\omega}$ be the environment obtained from ω by setting $\kappa = 0$. We use the coupling introduced in Section 2 and denote by $(\bar{X}_n)_n$ the random walk in the environment $\bar{\omega}$. It was shown in [5], that there exists a $\lambda_u < \infty$, such that for $\lambda > \lambda_u$, $\mathbb{E}_\lambda[\bar{T}_A] = \infty$, when $\bar{T}_A$ is the time $\bar{X}_n$ spends in A. In the following, fix such a λ. We claim that

$$\lim_{\kappa\to 0} \mathbb{E}_\lambda[T_A] = \infty. \tag{4.31}$$

Indeed, as in (4.2),

$$P_{\omega,\lambda}(X_n - X_{n-1} \neq \bar{X}_n - \bar{X}_{n-1} | X_{n-1} = \bar{X}_{n-1} = x) \leq C\kappa \tag{4.32}$$

for all $n \geq 1$ and $x \in \mathbb{Z}^2$. Let

$$D = \inf\{n \geq 1 : X_n - X_{n-1} \neq \bar{X}_n - \bar{X}_{n-1}\}.$$

Since (4.32) holds independent of x, D can be coupled with a geometric distributed random variable G with mean $(C\kappa)^{-1}$ independent of T_A such that $D \geq G$. Therefore,

$$\mathbb{E}_\lambda[T_A] \geq \mathbb{E}_\lambda[\bar{T}_A \wedge D] \geq \mathbb{E}_\lambda[\bar{T}_A \wedge G] \xrightarrow[\kappa\to 0]{} \mathbb{E}_\lambda[\bar{T}_A] = \infty.$$

Next, we define a sequence of ladder times $L_0, L_1, \dots$ with $L_0 = 0$ and let A_0 be the dead end starting at the origin (possibly empty). Inductively, let L_{i+1} be the first fresh epoch with $X_{L_{i+1}} \cdot e_1 > X_{L_i} \cdot e_1 + d(A_i)$ and let A_{i+1} be the dead end beginning at $X_{L_{i+1}}$. Since X_n is transient to the right, there are infinitely many ladder times. Note that $L_{i+1} - L_i \geq T_{A_i}$ and the random variables $T_{A_1}, T_{A_2}, \dots$ are iid under $\mathbb{P}_\lambda$ and satisfy (4.31). Additionally, the random variables $X_{L_{i+1}} \cdot e_1 - X_{L_i} \cdot e_1 = d(A_i) + 1$ are iid and have exponential moments (independent of κ) by Lemma 4.3. This implies for the speed

$$v_1(\lambda) = \lim_{n\to\infty} \frac{X_{L_n} \cdot e_1}{L_n} \leq \lim_{n\to\infty} \frac{\sum_{i=0}^n d(A_i) + 1}{\sum_{i=0}^n T_{A_i}} \leq \frac{C}{\mathbb{E}_\lambda[T_A]}.$$

Letting $\kappa \to 0$, we obtain (4.22) by (4.31). This completes the proof of Theorem 1.2.

Acknowledgements

We thank Andrew Barbour for helpful discussions about the proof of Lemma 4.1 and Alexander Telgen for discussions about the proof of Fact 2. Support of DFG (grant GA 582/8-1) is gratefully acknowledged.

References

[1] E. Aïdékon. Speed of the biased random walk on a Galton–Watson tree. *Probab. Theory Related Fields* **159** (3–4) (2014) 597–617. MR3230003

[2] M. Barma and D. Dhar. Directed diffusion in a percolation network. *J. Phys. C, Solid State Phys.* **16** (8) (1983) 1451.

[3] G. Ben Arous, A. Fribergh, N. Gantert and A. Hammond. Biased random walks on Galton–Watson trees with leaves. *Ann. Probab.* **40** (1) (2012) 280–338. MR2917774

[4] G. Ben Arous, A. Fribergh and V. Sidoravicius. Lyons–Pemantle–Peres monotonicity problem for high biases. *Comm. Pure Appl. Math.* **67** (4) (2014) 519–530. MR3168120

[5] N. Berger, N. Gantert and Y. Peres. The speed of biased random walk on percolation clusters. *Probab. Theory Related Fields* **126** (2) (2003) 221–242. MR1990055

[6] P. Billingsley. The invariance principle for dependent random variables. *Trans. Amer. Math. Soc.* **83** (1956) 250–268. MR0090923

[7] A. De Masi, P. A. Ferrari, S. Goldstein and W. D. Wick. An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Stat. Phys.* **55** (1988) 787–855. MR1003538

[8] D. Dhar. Diffusion and drift on percolation networks in an external field. *J. Phys. A: Math. Gen.* **17** (5) (1984) L257.

[9] P. G. Doyle and J. L. Snell. *Random Walks and Electric Networks*. Mathematical Association of America, Washington, DC, 1984. MR0920811

[10] A. Fribergh. Biased random walk in positive random conductances on $\mathbb{Z}^d$. *Ann. Probab.* **41** (6) (2013) 3910–3972. MR3161466

[11] A. Fribergh and A. Hammond. Phase transition for the speed of the biased random walk on the supercritical percolation cluster. *Comm. Pure Appl. Math.* **67** (2) (2014) 173–245. MR3149843

[12] N. Gantert, X. Guo and J. Nagel. Einstein relation and steady states for the random conductance model. *Ann. Probab.* **45** (4) (2017) 2533–2567. MR3693969

[13] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** (1986) 425–487. MR0871905

[14] T. Komorowski, C. Landim and S. Olla. *Fluctuations in Markov Processes: Time Symmetry and Martingale Approximation*, **345**. Springer, Berlin, 2012. MR2952852

[15] J. Lebowitz and H. Rost. The Einstein relation for the displacement of a test particle in a random environment. *Stochastic Process. Appl.* **54** (1994) 183–196. MR1307334

[16] R. Lyons, R. Pemantle and Y. Peres. Biased random walks on Galton–Watson trees. *Probab. Theory Related Fields* **106** (2) (1996) 249–264. MR1410689

[17] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Cambridge University Press, New York, 2016. Available at http://pages.iu.edu/~rdlyons/.

[18] L. Shen. Asymptotic properties of certain anisotropic walks in random media. *Ann. Appl. Probab.* **12** (2) (2002) 477–510. MR1910636

[19] A.-S. Sznitman. On the anisotropic walk on the supercritical percolation cluster. *Comm. Math. Phys.* **240** (1–2) (2003) 123–148. MR2004982

[20] A.-S. Sznitman and M. Zerner. A law of large numbers for random walks in random environment. *Ann. Probab.* **27** (4) (1999) 1851–1869. MR1742891