The Annals of Probability
2010, Vol. 38, No. 1, 234–276
DOI: 10.1214/09-AOP481

INVARIANCE PRINCIPLE FOR THE RANDOM CONDUCTANCE MODEL WITH UNBOUNDED CONDUCTANCES

BY M. T. BARLOW[1] AND J.-D. DEUSCHEL[2]

University of British Columbia and Technische Universität Berlin

We study a continuous time random walk X in an environment of i.i.d. random conductances $\mu_e \in [1, \infty)$. We obtain heat kernel bounds and prove a quenched invariance principle for X. This holds even when $\mathbb{E}\mu_e = \infty$.

1. Introduction. We consider the Euclidean lattice $\mathbb{Z}^d$ with $d \geq 2$. Let E_d, the set of nonoriented nearest neighbour bonds, and, writing $e = \{x, y\} \in E_d$, let $(\mu_e, e \in E_d)$ be nonnegative r.v., defined on a probability space $(\Omega, \mathbb{P})$. Throughout this paper we will assume that (μ_e) is stationary and ergodic, and that its law is invariant under symmetries of $\mathbb{Z}^d$. We write $\mu_{xy} = \mu_{\{x,y\}} = \mu_{yx}$, and let $\mu_{xy} = 0$ if $x \not\sim y$. Set

$$\mu_x = \sum_y \mu_{xy}, \qquad P(x, y) = \frac{\mu_{xy}}{\mu_x}. \tag{1.1}$$

There are two natural continuous time random walks associated with μ. Both jump according to the transitions $P(x, y)$. The first (the *constant speed random walk* or *CSRW*) waits at x for an exponential time with mean 1 while the second (the *variable speed random walk* or *VSRW*) waits at x for an exponential time with mean $1/\mu_x$. Write $\mathcal{L}_C$ and $\mathcal{L}_V$ for their generators, given by

$$\mathcal{L}_C f(x) = \mu_x^{-1} \sum_y \mu_{xy}\big(f(y) - f(x)\big), \tag{1.2}$$

$$\mathcal{L}_V f(x) = \sum_y \mu_{xy}\big(f(y) - f(x)\big). \tag{1.3}$$

Set

$$\mathcal{E}(f, g) = \frac{1}{2} \sum_{x \in \mathbb{Z}^d} \sum_{y \in \mathbb{Z}^d} \mu_{xy}\big(f(x) - f(y)\big)\big(g(x) - g(y)\big).$$

Received October 2008.
[1]Supported in part by NSERC (Canada) and EPSRC (UK).
[2]Supported in part by DFG-Forschergruppe 718.
AMS 2000 subject classifications. 60K37, 60F17, 82C41.
Key words and phrases. Random conductance model, heat kernel, invariance principle, ergodic, corrector.

Let $\nu_x = 1$, $x \in \mathbb{Z}^d$. It is easy to check that if f, g have finite support, then

$$\mathcal{E}(f,g) = -\sum_x g(x) \sum_y \mu_{xy}\big(f(y) - f(x)\big), \tag{1.4}$$

and so

$$\mathcal{E}(f,g) = -\langle \mathcal{L}_V f, g\rangle_\nu = -\langle \mathcal{L}_C f, g\rangle_\mu. \tag{1.5}$$

Thus the VSRW is the Markov process associated with the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on $L^2(\nu)$ and has stationary measure ν while the CSRW is the Markov process associated with the Dirichlet form $(\mathcal{E}, \mathcal{D}(\mathcal{E}))$ on $L^2(\mu)$ and has stationary measure μ.

Let $X = (X_t, t \geq 0, P_\omega^x, x \in \mathbb{Z}^d)$ be either the CSRW or the VSRW. Write $\mathcal{L}$ for its generator, θ for its invariant measure (so either $\theta = \nu$ or $\theta = \mu$) and let

$$q_t^\omega(x,y) = \frac{P_\omega^x(X_t = y)}{\theta_y} \tag{1.6}$$

be the transition density of X (or heat kernel associated with $\mathcal{L}$). This model, of a reversible (or symmetric) random walk in a random environment, is often called the *random conductance model* or RCM, particularly in the special case when (μ_e) are i.i.d. We are interested in the long-range behavior of X and, in particular, in obtaining heat kernel bounds for $q_t^\omega(x,y)$ and a quenched or $\mathbb{P}$-a.s. invariance principle for X. When $\mathbb{E}\mu_e < \infty$, an averaged invariance principle is obtained in [17].

We begin by discussing the case when (μ_e) are i.i.d. If $\mu_e = 0$ then X never jumps across e. So if $p_+ = \mathbb{P}(\mu_e > 0)$ is less than $p_c(d)$, the critical probability for bond percolation in $\mathbb{Z}^d$, then X is $\mathbb{P} \times P_\omega^x$-a.s. confined to a finite set. Thus we restrict to the case $p_+ > p_c$. A number of different authors have studied this model under various restrictions on the support of μ_e. If $\mu_e \in \{0, 1\}$ then this problem reduces to that of a random walk on (supercritical) percolation clusters (see [1] for heat kernel bounds, and [10, 29, 34] for quenched invariance principles). More generally it is useful to consider the following special cases:

Case 0. $c^{-1} \leq \mu_e \leq c$ for some $c \geq 1$;

Case 1. $0 \leq \mu_e \leq 1$;

Case 2. $1 \leq \mu_e < \infty$.

For case 0, heat kernel bounds follow from the results in [18, 19], and a quenched invariance principle is proved in [34]. Case 1 is treated in [11, 12, 30]. (The papers [11, 12] consider a discrete time random walk.) These papers prove an invariance principle, with a strictly positive diffusion constant σ^2. Further, [11] shows that Gaussian heat kernel bounds do not hold in general in this case.

In this paper we will look at case 2. There is not a great difference between the CSRW and VSRW in case 1, but in case 2, and in particular when $\mathbb{E}\mu_e = \infty$, the VSRW and CRSW do have different behaviors. Also, while the discrete time random walk with jump probabilities $P(x,y)$ given by (1.1) behaves in a similar fashion to the CSRW, there is no simple discrete time analogue of the VSRW in case 2. We remark that our result for the CSRW also gives an invariance principle for the discrete time random walk with jump probabilities $P(x, y)$.

Let

$$X_t^{(\varepsilon)} = \varepsilon X_{t/\varepsilon^2}, \qquad t \geq 0. \tag{1.7}$$

Our first main result is the following quenched functional central limit theorem (QFCLT):

THEOREM 1.1. *Let $d \geq 2$. Suppose that (μ_e) are i.i.d., and $\mu_e \geq 1$ $\mathbb{P}$-a.s.*

(a) *Let X be the VSRW. Then $\mathbb{P}$-a.s. $X^{(\varepsilon)}$ converges (under P_ω^0) in law to a Brownian motion on $\mathbb{R}^d$ with covariance matrix $\sigma_V^2 I$ where $\sigma_V > 0$ is nonrandom.*

(b) *Let X be the CSRW. Then $\mathbb{P}$-a.s. $X^{(\varepsilon)}$ converges (under P_ω^0) in law to a Brownian motion on $\mathbb{R}^d$ with covariance matrix $\sigma_C^2 I$ where*

$$\sigma_C^2 = \begin{cases} \sigma_V^2/(2d\mathbb{E}\mu_e), & \text{if } \mathbb{E}\mu_e < \infty, \\ 0, & \text{if } \mathbb{E}\mu_e = \infty. \end{cases}$$

We also have heat kernel bounds for the VSRW.

THEOREM 1.2. *Let $d \geq 2$. Suppose that (μ_e) are i.i.d. and $\mu_e \geq 1$ $\mathbb{P}$-a.s. Let $q_t^\omega(x, y)$ be the heat kernel for the VSRW. Let $\eta \in (0, 1)$. There exist r.v. U_x, $x \in \mathbb{Z}^d$, such that*

$$\mathbb{P}(U_x(\omega) \geq n) \leq c_1 \exp(-c_2 n^\eta) \tag{1.8}$$

and constants c_i (depending on d and the distribution of μ_e) such that the following hold.

(a) *For all x, y, t*

$$q_t^\omega(x, y) \leq c_3 t^{-d/2}. \tag{1.9}$$

(b) *If $|x - y| \vee t^{1/2} \geq U_x$, then*

$$q_t^\omega(x, y) \leq c_3 t^{-d/2} e^{-c_4 |x-y|^2/t} \qquad \text{when } t \geq |x - y|, \tag{1.10}$$

$$q_t^\omega(x, y) \leq c_3 \exp\big(-c_4 |x - y| \big(1 \vee \log(|x - y|/t)\big)\big) \qquad \text{when } t \leq |x - y|. \tag{1.11}$$

(c) *Let $x, y \in \mathbb{Z}^d$, $t > 0$. Then*

$$q_t^\omega(x, y) \geq c_5 t^{-d/2} e^{-c_6 |x-y|^2/t} \qquad \text{if } t \geq U_x^2 \vee |x - y|^{1+\eta}. \tag{1.12}$$

(d) *Let $x, y \in \mathbb{Z}^d$ and $t \geq c_7 \vee |x - y|^{1+\eta}$. Then*

$$c_8 t^{-d/2} e^{-c_9 |x-y|^2/t} \leq \mathbb{E} q_t^\omega(x, y) \leq c_{10} t^{-d/2} e^{-c_{11} |x-y|^2/t}. \tag{1.13}$$

Using Theorems 1.1 and 1.2 we can obtain a parabolic Harnack inequality (PHI) for q_t^ω (for the VSRW) by the same arguments as in [3] (see Theorem 4.7). Since the CSRW is a time change of the VSRW, harmonic functions and Green's functions are the same for both processes. The PHI for the VSRW implies an elliptic Harnack inequality (see Corollary 4.8) which therefore holds for both CSRW and VSRW. Combining the invariance principle and the PHI, we obtain, using the methods of [3], a local limit theorem for the VSRW (see Theorem 5.14).

When $d \geq 3$ the calculations in Section 6 of [3] then give bounds on the Green's function $g_\omega(x, y)$ defined by

$$g_\omega(x, y) = \int_0^\infty q_t^\omega(x, y)\, dt. \tag{1.14}$$

THEOREM 1.3. *Let $d \geq 3$, and suppose that (μ_e) are i.i.d. and $\mu_e \geq 1$ $\mathbb{P}$-a.s.*

(a) *There exist constants $c_1, \dots, c_4$ and r.v. U_x, $x \in \mathbb{Z}^d$ such that*

$$\mathbb{P}(U_x \geq n) \leq c_1 \exp(-c_2 n^{1/3}) \tag{1.15}$$

and

$$\frac{c_3}{|x-y|^{d-2}} \leq g_\omega(x, y) \leq \frac{c_4}{|x-y|^{d-2}} \qquad \text{if } |x-y| \geq U_x \wedge U_y. \tag{1.16}$$

(b) *Let $C = \Gamma(\frac{d}{2} - 1)/(2\pi^{d/2}\sigma_V^2)$. For any $\varepsilon > 0$ there exists $M = M(\varepsilon, \omega)$ with $\mathbb{P}(M < \infty) = 1$ such that*

$$\frac{(1-\varepsilon)C}{|x|^{d-2}} \leq g_\omega(0, x) \leq \frac{(1+\varepsilon)C}{|x|^{d-2}} \qquad \text{for } |x| > M(\omega). \tag{1.17}$$

(c) *We have*, $\mathbb{P}$-*a.s.*,

$$\lim_{|x|\to\infty} |x|^{2-d} g_\omega(0, x) = \lim_{|x|\to\infty} |x|^{2-d} \mathbb{E} g_\omega(0, x) = C. \tag{1.18}$$

REMARK 1.4. (b) and (c) in Theorem 1.3 use the QFCLT, which in turn uses the ergodic theorem. As we do not have any rate of convergence in the QFCLT this means that [unlike the r.v. U_x in (a)] we have no control on the tail of the r.v. M in (b).

The main difference between the RCM and the percolation case is the possibility of traps. Suppose $e = \{x, y\}$ is a bond with $\mu_e = K \gg 1$, and that all the other bonds e' adjacent to x and y have $\mu_{e'} \simeq 1$. Then $P(x, y) \simeq 1 - c/K$, so X will jump between x and y $O(K)$ times before leaving $\{x, y\}$, and thus the CSRW will be trapped for a time of order K in $\{x, y\}$. However, for the VSRW each jump takes only $O(K^{-1})$, so the total time spent in $\{x, y\}$ is only $O(1)$. A similar effect will arise from finite clusters of bonds of high conductivity.

The presence of traps of this kind is why we have, when $\mathbb{E}\mu_e = \infty$, that the diffusion constant σ_C^2 for the CSRW is zero. In this case it is natural to ask if a different scaling will give a nontrivial limit. There is a connection here with "aging" (see [6–8]), and in [2] it is proved that if the tail distribution $\mathbb{P}(\mu_e > t) \sim t^{-\alpha}$, then $\varepsilon^\alpha X_{t/\varepsilon^2}$ converges to the "fractional kinetic" motion with parameter α (see [8]).

While we have written this paper for the case of i.i.d. conductances μ_e, our arguments do not require the full strength of this. In case 0, when the conductances are bounded (both above and below), then uniform upper and lower Gaussian heat kernel estimates, as in Theorem 1.2, are well known (see [18]). It follows (see Remark 6.3) that Theorem 1.1. holds for any stationary ergodic environment. On the other hand, in the unbounded case 2, there exist stationary ergodic environments such that the VSRW can explode in finite time; for an example, see Remark 6.6 below.

For the Gaussian bounds in Theorem 1.2 we need to control the sizes of the clusters of high conductivity which is done by comparison between the graph metric $d(x,y)$ and a new metric $\widetilde{d}$ (introduced by Davies in [15, 16]) which is adapted to the structure of the random walk and satisfies $\widetilde{d}(x,y) \le \mu_{xy}^{-1/2}$ when $x \sim y$. This new metric is constructed by a first passage percolation procedure, and in this paper we have used first passage percolation arguments to compare the two metrics (see Lemma 4.2). These arguments use estimates from [25] which in turn use the fact that μ_e are i.i.d. However, we could also have used a direct argument as in [12], Lemma 3.1 or [30], Lemma 5.3. Once we have the Gaussian bounds (with sufficiently good control on the tails of the r.v. U_x in Theorem 1.2), the quenched invariance principle follows with no further hypotheses on $\{\mu_e, e \in E_d\}$ other than that it is stationary, symmetric and ergodic. Theorem 6.1 summarizes the general situation.

The structure of this paper is as follows. In Sections 2 and 3 we study a deterministic graph $\Gamma = (G, E)$ with edge weights μ_{xy}. Under certain conditions (which are $\mathbb{P}$-a.s. satisfied by the VSRW on the i.i.d. RCM) we obtain heat kernel bounds in this setup. Our approach uses similar methods to those used in [1] for percolation clusters. However, in [1] the Carne–Varopoulos "long-range" bounds played an essential role at various points. These bounds do not hold for the VSRW, and instead we use more general upper bounds obtained by Davies [15, 16], which are in terms of the metric $\widetilde{d}(x,y)$. The same metric $\widetilde{d}$ is also needed to control $P_\omega^x(\widetilde{d}(X_t, x) \ge \lambda t^{1/2})$ which is the key step in obtaining general Gaussian upper bounds. Similar bounds, in the context of weighted Laplacians on manifolds, are obtained by Grigor'yan [24]; here the metric $\widetilde{d}$ is the Riemannian metric. Very recently, and independently, Mourrat [31] has obtained upper bounds for the VSRW which in certain cases improve on those in Theorem 1.2.

Once one has upper bounds, lower bounds follow by the same arguments as in [1], Section 5 (see Section 3). In Section 4 we then prove Theorem 1.2.

Section 5 proves the invariance principle. We begin with the VSRW. The basic technique in the proof (as in many previous papers such as [10, 12, 17, 27–29]) is to associate with X_t a process Z_t on $\Omega = [1, \infty]^{\mathbb{E}^d}$ which is the environment seen from the random walk. More precisely, for each $x \in \mathbb{Z}^d$, let $T_x : \Omega \to \Omega$ be given by

$$T_x(\omega)(z, w) = \omega(z + x, w + x).$$

Assuming that $X_0 = 0$ we define

$$Z_t(\omega) = T_{X_t(\omega)}(\omega). \tag{1.19}$$

One seeks to use the process Z to construct the "corrector," that is, a map $\chi : \Omega \times \mathbb{Z}^d \to \mathbb{R}^d$ such that

$$M_t(\omega) = X_t(\omega) - \chi(\omega, X_t(\omega)), \qquad t \geq 0, \tag{1.20}$$

is a P_ω^0-martingale. Once one has constructed the corrector, showing the invariance principle for the rescaled martingale $\varepsilon M_{t/\varepsilon^2}$ is standard, and using results from [12, 34], the heat kernel estimates in Theorem 1.2, together with the sublinear growth of χ, imply that

$$\lim_{\varepsilon \to 0} \varepsilon \chi(\omega, X_{t/\varepsilon^2}) = 0 \qquad \text{in } P_\omega^0\text{-probability}. \tag{1.21}$$

However, the standard construction of the corrector is based on $L^2(\mathbb{P})$ calculus, which requires finiteness of the first moment of the conductance (see [17], page 816). In our case we wish to handle the case when $\mathbb{E}\mu_e = \infty$, and so we need an alternative approach. (We remark that if $\mathbb{E}\mu_e = \infty$ then it is not easy to find suitable function spaces on Ω which give a core for the Dirichlet form associated with Z.)

Our solution relies on discretization. We define $\widehat{X}_n = X_n$, $n \in \mathbb{Z}_+$, and consider the process

$$\widehat{X}_t^{(\varepsilon)} = \varepsilon \widehat{X}_{\lfloor t/\varepsilon^2 \rfloor}. \tag{1.22}$$

We can control $\sup_{t \leq T} |X_t^{(\varepsilon)} - \widehat{X}_t^{(\varepsilon)}|$ (see Lemma 4.12), so an invariance principle for $X^{(\varepsilon)}$ will follow from one for $\widehat{X}^{(\varepsilon)}$. The process $\widehat{X}$ does not have bounded jumps—in fact it jumps anywhere in $\mathbb{Z}^d$ with positive probability. However, the long-range bounds on $q_t^\omega(x, y)$ in (1.11) give good control on these jumps, and, in particular, the bounds on $q_1^\omega(x, y)$ imply that

$$\mathbb{E} E_\omega^0 |X_1|^2 < \infty, \tag{1.23}$$

which is the key L^2 condition on $\widehat{X}$ for the construction of the corrector. As we will see in Section 5, looking at the discrete time process does actually introduce some simplifications in the construction of the corrector χ. In the end (see Remark 5.15) it will turn out that the "discrete time" corrector χ also satisfies (1.20).

Finally, a short Section 6 makes some remarks on more general environments, and gives an example (a one-sided spanning tree) where the process X fails to be conservative, and so the invariance principle fails.

2. Transition density upper bounds on a fixed graph. Let $\Gamma = (G, E)$ be an infinite (deterministic) graph, $\mu_e, e \in E$, be bond conductances and ν be a measure on G. We make the following assumptions on (G, E), μ and ν.

ASSUMPTION 2.1. (1) Γ is connected.
(2) The vertex degree is uniformly bounded,

$$|\{y : y \sim x\}| \le C_D \qquad \text{for all } x \in G. \tag{2.1}$$

(3) $\mu_e > 0$ for all $e \in E$.
(4) There exists $C_M \ge 1$ such that

$$C_M^{-1} \le \nu_x = \nu(\{x\}) \le C_M \qquad \text{for all } x \in G. \tag{2.2}$$

The results of this section do not explicitly require a strictly positive lower bound on μ_e; however, a later assumption [see Assumption 2.6(2)] will impose some control on the edges e with μ_e small.

We write μ_{xy} for $\mu_{\{x,y\}}$, and set $\mu_{xy} = 0$ if $x \not\sim y$. Let $d(x, y)$ be the usual graph distance on G, and write

$$B(x,r) = \{y : d(x, y) < r\}. \tag{2.3}$$

Let $C_A < \infty$. We now construct, by a first passage percolation procedure, a second metric $\widetilde{d}$ on G satisfying

$$(C_A^{-2} \vee \mu_{yy'})|\widetilde{d}(x, y) - \widetilde{d}(x, y')|^2 \le 1 \qquad \text{for every } x \in G, y \sim y'. \tag{2.4}$$

We write $\widetilde{B}(x,r) = \{y : \widetilde{d}(x, y) < r\}$ for balls in the metric $\widetilde{d}$. (In this paper we can take $C_A = 1$, but for possible future extensions we treat the general case.) To construct $\widetilde{d}$ we assign waiting times

$$t(e) = C_A \wedge \mu_e^{-1/2}, \qquad e \in E, \tag{2.5}$$

and then take $\widetilde{d}(x, y)$ to be the shortest journey time between x and y. More precisely,

$$\widetilde{d}(x, y) = \inf\left\{\sum_{i=1}^{n} t(e_i)\right\}, \tag{2.6}$$

where the infimum is taken over paths $(e_1, \dots, e_n)$ from x to y. Since we do not have a strictly positive uniform lower bound on $t(e)$, in general there may not be a minimizing path. However, such paths will a.s. exist when $t(e_i)$ are i.i.d. positive random variables.

LEMMA 2.2. *The metric $\widetilde{d}$ constructed above satisfies* (2.4).

PROOF. Let $e = \{y, z\}$; then $\widetilde{d}(x, z) \le \widetilde{d}(x, y) + t(e)$, and using (2.5) gives (2.4). $\square$

Recall that

$$\mathcal{E}(f, g) = \frac{1}{2} \sum_x \sum_{y \sim x} \mu_{xy} \big(f(y) - f(x)\big)\big(g(y) - g(x)\big).$$

Let $\mu_x = \sum_{y \sim x} \mu_{xy}$, and extend μ to a measure on G. Then $\mathcal{E}(f, g)$ is defined for $f, g \in L^2(G, \mu)$.

Let $X = (X_t, t \in [0, \infty), P^x, x \in G)$ be the continuous time Markov chain on G with generator

$$\mathcal{L} f(x) = \nu_x^{-1} \sum_y \mu_{xy} \big(f(y) - f(x)\big).$$

At this point we cannot exclude the possibility that X may explode, and we write ζ for the lifetime of X. Let

$$\|f\|_{\mathcal{E}_1}^2 = \mathcal{E}(f, f) + \|f\|_{L^2(\nu)}^2$$

and $\mathcal{F}$ be the closure of the set of functions on G with finite support with respect to $\|f\|_{\mathcal{E}_1}$. Then X is the Markov process associated with the Dirichlet form $(\mathcal{E}, \mathcal{F})$ on $L^2(G, \nu)$ (see [22]). Let $q_t(x, y)$ be the transition density (heat kernel) of X with respect to ν:

$$q_t(x, y) = \frac{P^x(X_t = y)}{\nu_y}.$$

We begin by using the results of Davies [15, 16] to obtain long-range bounds on q_t. By Proposition 5 of [15], we have

$$q_t(x, y) \le (\nu_x \nu_y)^{-1/2} \inf_{\psi \in L^\infty(G)} \exp\big(\psi(y) - \psi(x) + \Lambda(\psi) t\big), \tag{2.7}$$

where $\Lambda(\psi) = \sup_x b(\psi, x)$, and

$$b(\psi, x) = \frac{1}{2\nu_x} \sum_{y \sim x} \mu_{xy} \big(e^{\psi(x) - \psi(y)} + e^{\psi(y) - \psi(x)} - 2\big). \tag{2.8}$$

THEOREM 2.3. *Assume (G, E) and μ satisfy Assumption 2.1. There exist constants $c_1, \dots, c_4$ (depending only on C_A, C_D, C_M) such that the following hold.*

(a) *If $x, y \in G$ and $\widetilde{D} = \widetilde{d}(x, y) \le c_1 t$, then*

$$q_t(x, y) \le c_2 \exp(-c_3 \widetilde{D}^2 / t). \tag{2.9}$$

(b) *If $x, y \in G$ and $\widetilde{D} = \widetilde{d}(x, y) \ge c_1 t$, then*

$$q_t(x, y) \le c_2 \exp\big(-c_4 \widetilde{D}\big(1 \vee \log(\widetilde{D}/t)\big)\big). \tag{2.10}$$

PROOF. Fix $x_0, y_0 \in G$, let $t > 0$, and write $\widetilde{D} = \widetilde{d}(x_0, y_0)$. Let $\lambda > 0$ and set

$$\psi_\lambda(x) = -\lambda(\widetilde{D} \wedge \widetilde{d}(x_0, x)), \qquad b(\lambda) = \sup_x b(\psi_\lambda, x).$$

Let $x \in G$, $y \sim x$ and write $\widetilde{\mu}_{xy} = C_A^{-2} \vee \mu_{xy}$,

$$J_{xy} = \mu_{xy}(e^{\psi_\lambda(x)-\psi_\lambda(y)} + e^{\psi_\lambda(y)-\psi_\lambda(x)} - 2).$$

Then as $|\psi_\lambda(x) - \psi_\lambda(y)| \le \lambda \widetilde{\mu}_{xy}^{-1/2}$, and $\cosh(x)$ is increasing on $[0, \infty)$,

$$J_{xy} \le 2\mu_{xy}(\cosh(\lambda\widetilde{\mu}_{xy}^{-1/2}) - 1) \le 2\widetilde{\mu}_{xy}(\cosh(\lambda\widetilde{\mu}_{xy}^{-1/2}) - 1). \tag{2.11}$$

Using the power series for cosh we have that the right-hand side of (2.11) is decreasing in $\widetilde{\mu}_{xy}$, so

$$J_{xy} \le C_A^{-2}(e^{C_A\lambda} + e^{-C_A\lambda} - 2).$$

Hence

$$b(\psi_\lambda, x) \le \tfrac{1}{2} C_M C_D C_A^{-2}(e^{C_A\lambda} + e^{-C_A\lambda} - 2).$$

Let $f(x) = e^x + e^{-x} - 2$; then $b(\lambda) \le c_7 f(C_A\lambda)$. Thus by (2.7), and writing $y = C_A\lambda$,

$$\begin{aligned} q_t(x_0, y_0) &\le C_M \inf_\lambda \exp(-\lambda\widetilde{D} + c_7 t f(C_A\lambda)) \\ &\le C_M \exp\biggl(\frac{\widetilde{D}}{C_A}\biggl(\inf_{y>0}\biggl(-y + \frac{C_A c_7 t}{\widetilde{D}} f(y)\biggr)\biggr)\biggr). \end{aligned} \tag{2.12}$$

So if

$$F(s) = \inf_{y>0}(-y + (2s)^{-1}(e^y + e^{-y} - 2)),$$

then

$$q_t(x_0, y_0) \le C_M \exp\biggl(\frac{\widetilde{D}}{C_A} F\biggl(\frac{\widetilde{D}}{2C_A c_7 t}\biggr)\biggr) \tag{2.13}$$

and it remains to bound F.

We have (see [15], page 70) that

$$F(s) = s^{-1}((1 + s^s)^{1/2} - 1) - \log(s + (1 + s^2)^{1/2})$$

and also that $F(s) \le -s/2(1 - s^2/10)$ for $s > 0$. Hence, if $s \le 3$, then $F(s) \le -s/20$ while if $s \ge e$, then

$$F(s) \le 1 - \log(2s) = -\log(2s/e).$$

Substituting in (2.13) completes the proof. □

REMARK 2.4. Note that if $\widetilde{D} = ct$ then both (2.9) and (2.10) give a bound of the form $ce^{-t/c}$.

Since μ_e are not bounded above, the process X may explode. The following condition is enough to exclude this.

LEMMA 2.5. *Suppose there exists $x \in G$ and $\theta > 0$ such that*

$$\sum_{y\in G} \exp(-\theta \widetilde{d}(x,y))\nu_y < \infty. \tag{2.14}$$

Then X is conservative.

PROOF. Let ζ be the lifetime of X. Then as Γ is connected it is easy to see that either $P^y(\zeta = \infty) = 1$ for all $y \in G$, or else $P^y(\zeta < t) > 0$ for all $y \in G$, $t > 0$.

For $n \geq C_A^{-2}$ let $\mu_{xy}^{(n)} = n \wedge \mu_{xy}$, $X^{(n)}$ be the process associated with the conductances $\mu^{(n)}$, and $q_t^{(n)}(x,y)$ be the transition density of $X^{(n)}$ with respect to ν. We have $q_t(x,y) = \lim_{n\to\infty} q_t^{(n)}(x,y)$. Note that each $X^{(n)}$ is conservative, and that the bounds in Theorem 2.3 hold (with the same constants c_i) for each $q^{(n)}$. With (2.2) the condition (2.14) implies that $\widetilde{B}(x,R)$ is finite for each $R > 0$.

Let $t > 0$. With constants c_i as in Theorem 2.3, choose r large enough so that $r > c_1 t$ and $c_4(1 \vee \log(r/t)) \geq \theta$. So, if $R \geq r$, using (2.10),

$$\sum_{y\in \widetilde{B}(x,R)^c} q_t^{(n)}(x,y)\nu_y \leq \sum_{y\in \widetilde{B}(x,R)^c} c_2 \exp(-\theta \widetilde{d}(x,y))\nu_y < \infty. \tag{2.15}$$

Let $\varepsilon > 0$; then we can take R large enough so that the right-hand side of (2.15) is less than ε. Thus, as $X^{(n)}$ is conservative, for all n,

$$\sum_{y\in \widetilde{B}(x,R)} q_t^{(n)}(x,y)\nu_y > 1 - \varepsilon.$$

So,

$$P^x(\zeta > t) \geq \sum_{y\in \widetilde{B}(x,R)} q_t(x,y)\nu_y = \lim_{n\to\infty} \sum_{y\in \widetilde{B}(x,R)} q_t^{(n)}(x,y)\nu_y \geq 1 - \varepsilon.$$

Therefore $P^x(\zeta > t) = 1$ for all t, proving that X is conservative. $\square$

We now make further assumptions on the graph Γ and the conductances μ. As we will see in Section 4, it is easy to check these for the random conductance model on $\mathbb{Z}^d$ when $\mu_e \geq 1$.

ASSUMPTION 2.6. (1) There exists $d \geq 1$ and $C_V < \infty$ such that

$$\nu(B(x,r)) \leq C_V r^d \qquad \text{for all } x \in G, r \geq 1. \tag{2.16}$$

(2) There exists a constant C_N such that the following Nash inequality holds. If $f \in L^1(G,\nu) \cap L^2(G,\nu)$, then, writing $\|f\|_p$ for norms in $L^p(G,\nu)$,

$$\mathcal{E}(f,f) \geq C_N \|f\|_2^{2+4/d} \|f\|_1^{-4/d}. \tag{2.17}$$

REMARK 2.7. Note that (2) above does place some restrictions on the edges e with μ_e small. For example, taking $x \in G$ and $f = 1_x$, (2.17) gives

$$\sum_y \mu_{xy} \geq C_N \nu_x^{1-2/d}.$$

By [14] we have:

LEMMA 2.8. *Suppose* (2.17) *holds. Then*

$$q_t(x,y) \leq ct^{-d/2}, \qquad x, y \in G, t > 0. \tag{2.18}$$

To obtain better control of $q_t(x,y)$ when $d(x,y)$ is large we need to compare the metrics d and $\widetilde{d}$. Note first that $\widetilde{d}(x,y) \leq C_A \wedge \mu_{xy}^{-1/2} \leq C_A$ when $x \sim y$, so

$$\widetilde{d}(x,y) \leq C_A d(x,y), \qquad x, y \in G. \tag{2.19}$$

DEFINITION 2.9. Let $\lambda \geq 1$, $\eta \in (0,1)$. Let $x \in G$, $r \in [1,\infty)$. We say (x,r) is λ-*good* if $\widetilde{B}(x, n/\lambda) \subset B(x,n)$ for all $n \geq r$, $n \in \mathbb{N}$. We say (x, R_0) is λ-*very good* if for all $R \geq R_0$, (y,r) is λ-good for all $y \in B(x,R)$, $r \geq R^\eta$, $r \in \mathbb{N}$. Note that if (x,R_0) is λ-*very good* then (x,R_1) is λ-*very good* for all $R_1 \geq R_0$. For $x \in G$ let $V_x = V_x(\lambda)$ be the smallest integer such that (x, V_x) is λ-very good.

Note. Unlike the definitions in [1], the event that (x,r) is λ-good depends on the structure of Γ "at infinity."

LEMMA 2.10. *Suppose* (x,R) *is* λ-*good.*
(a) *If* $d(x,y) \geq R$,

$$\lambda^{-1} d(x,y) \leq \widetilde{d}(x,y) \leq C_A d(x,y). \tag{2.20}$$

(b) *If* $R' \geq (2R) \vee 2(1 + C_A\lambda) d(x,x')$, *then* (x',R') *is* 2λ-*good.*

PROOF. (a) The upper bound is given in (2.19). For the lower bound, let $s = d(x,y) \geq R$. Then $y \notin B(x,s)$, so $y \notin \widetilde{B}(x, s/\lambda)$ and thus $\widetilde{d}(x,y) \geq s/\lambda$.
(b) Let $\widetilde{r} = \widetilde{d}(x,x')$, $r = d(x,x')$, and $s \geq R'$. Then as $s/2 \geq R$,

$$\widetilde{B}(x', s/2\lambda) \subset \widetilde{B}(x, \widetilde{r} + s/2\lambda) \subset B(x, \lambda\widetilde{r} + s/2).$$

So, using (2.20), $\widetilde{B}(x', s/2\lambda) \subset B(x', (1+\lambda C_A)r + s/2) \subset B(x', s)$. □

LEMMA 2.11. *Let $x \in G$, $\theta > 0$, $r \geq 1$. If (x, r) is λ-good, then*

$$\sum_{y \in B(x,r)^c} \exp(-\theta \widetilde{d}(x, y))\nu_y \leq \begin{cases} c(\lambda) r^d e^{-cr\theta}, & \text{if } r\theta \geq 1, \\ c(\lambda)\theta^{-d}, & \text{if } r\theta < 1. \end{cases} \tag{2.21}$$

In particular, X is conservative.

PROOF. Write I for the left-hand side of (2.21), and $D_n = B(x, 2^n r) - B(x, 2^{n-1} r)$. Then

$$I \leq \sum_{n=1}^{\infty} \sum_{y \in D_n} \exp\bigl(-\theta d(x, y)/\lambda\bigr)\nu_y \leq \sum_{n=1}^{\infty} C_V (2^n r)^d \exp(-2^{n-1} r\theta/\lambda). \tag{2.22}$$

If $\alpha > 0$, $d \geq 1$ then there exists $c_1 = c_1(d)$ such that

$$\sum_{n=1}^{\infty} 2^{nd} e^{-\alpha 2^n} \leq \begin{cases} c_1 e^{-\alpha}, & \text{if } \alpha \geq 1, \\ c_1 \alpha^{-d}, & \text{if } \alpha < 1, \end{cases}$$

and using these bounds in (2.22) completes the proof. $\square$

LEMMA 2.12. *Let $x \in G$ and suppose that (x, r) is λ-good. If $t \in (0, 1)$, then*

$$E^x d(x, X_t)^p \leq c(\lambda) r^{p+d}. \tag{2.23}$$

PROOF. Using the bound (2.10) a similar calculation to that in Lemma 2.11 gives

$$\begin{aligned} E^x d(x, X_t)^p &\leq r^p + \sum_{n=1}^{\infty} C_V (2^n r)^{d+p} e^{-c\widetilde{d}(x,y)} \\ &\leq r^p + c r^{d+p} \sum_{n=1}^{\infty} 2^{n(d+p)} \exp(-c' 2^n r/\lambda) \leq c r^{d+p}. \end{aligned}$$

$\square$

We now follow the arguments in [1], Section 3 (the "Bass–Nash method") to obtain Gaussian upper bounds on $q_t(x, y)$. As in Lemma 2.5 for $1 \leq n \leq \infty$, let $\mu_{xy}^{(n)} = \mu_{xy} \wedge n$, $X^{(n)}$ be the associated VSRW, and $q^{(n)}(x, y)$ be the transition density of $X^{(n)}$. Let $x_0 \in G$, and set

$$M_n(t) = M_n(x_0, t) = E^{x_0} \widetilde{d}\bigl(x_0, X_t^{(n)}\bigr) = \sum_y \widetilde{d}(x_0, y) q_t^{(n)}(x_0, y)\nu_y, \tag{2.24}$$

$$Q_n(t) = Q_n(x_0, t) = -\sum_y q_t^{(n)}(x_0, y) \log q_t^{(n)}(x_0, y)\nu_y. \tag{2.25}$$

The following three inequalities lead, by Nash's argument [32], to upper bounds on $M_n(t)$ which are uniform in n. [We remark that we only need the approximations $X^{(n)}$ to justify an interchange of sums in part (c) of the following lemma.]

LEMMA 2.13. *Let $x_0 \in G$ and $r \geq 1$. Suppose (x_0, r) is λ-good, and $1 \leq n < \infty$. There exist constants c_i, independent of n, such that the following hold.*

(a) *We have, for $t > 0$,*

$$Q_n(x_0, t) \geq c_1 + \tfrac{1}{2} d \log t. \tag{2.26}$$

(b)

$$M_n(x_0, t) \geq c_2 e^{Q_n(x_0,t)/d} \qquad \textit{provided either } M_n(x_0, t) \geq r \textit{ or } t \geq c_3 r^2. \tag{2.27}$$

(c) *For $t > 0$,*

$$Q_n'(t) \geq c_4 M_n'(t)^2. \tag{2.28}$$

PROOF. We write $Q_n(t)$, $M_n(t)$ for $Q_n(x_0, t)$, $M_n(x_0, t)$. (a) is immediate from (2.18) and the fact that since $X^{(n)}$ is conservative, $\sum_y q_t^{(n)}(x_0, y)\nu_y = 1$.

(b) The proof is similar to those in [1, 4, 32]. First note that (2.16) and Lemma 2.8 give that

$$M_n(t) \geq r \qquad \text{provided } t \geq c_3 r^2. \tag{2.29}$$

By Lemma 2.11, provided $ar \leq 1$,

$$\begin{aligned}\sum_{y \in G} e^{-a\widetilde{d}(x_0,y)} \nu_y &\leq \sum_{y \in B(x_0,r)} e^{-a\widetilde{d}(x_0,y)} \nu_y + \sum_{y \notin B(x_0,r)} e^{-a\widetilde{d}(x_0,y)} \nu_y \\ &\leq c r^d + c a^{-d} \leq c a^{-d}.\end{aligned}$$

Now $u(\log u + \theta) \geq -e^{-1-\theta}$ for $u > 0$. So, setting $\theta = a\widetilde{d}(x_0, y) + b$, where $a \leq 1/r$,

$$\begin{aligned}-Q_n(t) + aM_n(t) + b &= \sum_{y \in G} q_t^{(n)}(x_0, y)\big(\log q_t^{(n)}(x_0, y) + a\widetilde{d}(x_0, y) + b\big)\nu_y \\ &\geq -\sum_{y \in G} e^{-1-a\widetilde{d}(x_0,y)-b} \nu_y \\ &\geq -e^{-1-b} \sum_{y \in G} e^{-a\widetilde{d}(x_0,y)} \nu_y \geq -c_5 e^{-b} a^{-d}.\end{aligned}$$

Setting $a = 1/M_n(t)$ and $e^b = M_n(t)^d = a^{-d}$, we obtain

$$-Q_n(t) + 1 + d \log M_n(t) \geq -c_6$$

and rearranging gives (b).

(c) Set $f_t(x) = q_t^{(n)}(x_0, x)$, and let $b_t(x, y) = f_t(x) + f_t(y)$. We have, using (2.4),

$$\begin{aligned}M_n'(t) &= \sum_y \widetilde{d}(x_0, y) \frac{\partial f_t(y)}{\partial t} \nu_y = \sum_y \widetilde{d}(x_0, y) \mathcal{L}^{(n)} f_t(y) \nu_y \\ &= -\frac{1}{2} \sum_x \sum_y \mu_{xy} \big(\widetilde{d}(x_0, y) - \widetilde{d}(x_0, x)\big)\big(f_t(y) - f_t(x)\big);\end{aligned}$$

the final interchange of sums can be justified using (2.9) and (2.10) and the fact that $\mu^{(n)}$ is uniformly bounded. So using (2.4),

$$
\begin{aligned}
M_n'(t) &\le \frac{1}{2}\sum_x\sum_{y\sim x}\bigl(\mu_{xy}^{1/2}|\widetilde{d}(x_0,y)-\widetilde{d}(x_0,x)|b_t(x,y)^{1/2}\bigr)\biggl(\mu_{xy}^{1/2}\frac{|f_t(y)-f_t(x)|}{b_t(x,y)^{1/2}}\biggr)\\
&\le \frac{1}{2}\sum_x\sum_{y\sim x}\bigl(b_t(x,y)^{1/2}\bigr)\biggl(\mu_{xy}^{1/2}\frac{|f_t(y)-f_t(x)|}{b_t(x,y)^{1/2}}\biggr)\\
&\le \frac{1}{2}\biggl(\sum_x\sum_{y\sim x}b_t(x,y)\biggr)^{1/2}\biggl(\sum_x\sum_y\mu_{xy}\frac{(f_t(y)-f_t(x))^2}{b_t(x,y)}\biggr)^{1/2}.
\end{aligned}
$$

Now

$$
\sum_x\sum_{y\sim x}b_t(x,y)=2\sum_x\sum_{y\sim x}f_t(x)\le 2C_DC_M\sum_x f_t(x)\nu_x=2C_DC_M. \tag{2.30}
$$

So,

$$
M_n'(t)^2\le c\sum_x\sum_y\mu_{xy}\frac{(f_t(y)-f_t(x))^2}{f_t(x)+f_t(y)}.
$$

Since we have, for $u,v>0$,

$$
\frac{(u-v)^2}{u+v}\le (u-v)(\log u-\log v),
$$

we deduce

$$
M_n'(t)^2\le c\sum_x\sum_y\mu_{xy}\bigl(f_t(y)-f_t(x)\bigr)\bigl(\log f_t(y)-\log f_t(x)\bigr).
$$

Thus

$$
\begin{aligned}
Q_n'(t)&=-\sum_y\bigl(1+\log f_t(y)\bigr)\mathcal{L}^{(n)}f_t(y)\nu_y\\
&=\frac{1}{2}\sum_x\sum_y\mu_{xy}\bigl(\log f_t(y)-\log f_t(x)\bigr)\bigl(f_t(y)-f_t(x)\bigr) \qquad (2.31)\\
&\ge cM_n'(t)^2;
\end{aligned}
$$

where again the interchange of sums uses (2.9) and (2.10) and the fact that $\mu^{(n)}$ is uniformly bounded. $\square$

PROPOSITION 2.14. *Let $x_0\in G$, $r\ge 1$ and (x_0,r) be λ-good. Then*

$$
c_1t^{1/2}\le M_\infty(x_0,t)\le c_2t^{1/2}\qquad \textit{for } t\ge c_3r^2. \tag{2.32}
$$

PROOF. Note that the lower bound is immediate from Lemma 2.8. For the upper bound let first $n < \infty$. Set $R_n(t) = d^{-1}(Q_n(t) - c_1 - \frac{1}{2}d\log t)$, so that $R_n(t) \geq 0$ by (2.26). Then

$$M_n(t) = \int_0^t M_n'(s)\,ds \leq c\int_0^t Q_n'(s)^{1/2}\,ds \leq c\int_0^t \left(R_n'(s) + \frac{1}{2s}\right)^{1/2} ds.$$

Using the inequality $(a+b)^{1/2} \leq b^{1/2} + a/(2b)^{1/2}$ and integrating by parts we obtain

$$\begin{aligned} M_n(t) &\leq ct^{1/2} + c\int_0^t s^{1/2} R_n'(s)\,ds \\ &\leq ct^{1/2} + c\bigl(1 + R_n(t)t^{1/2}\bigr) \leq ct^{1/2}\bigl(1 + R_n(t)\bigr). \end{aligned}$$

By (2.27) we also have $M_n(t) \geq t^{1/2}e^{R_n(t)}$ for $t > cr^2$. Thus $R_n(t)$ is bounded for $t > cr^2$ and this implies that

$$c_1 t^{1/2} \leq M_n(x_0, t) \leq c_2 t^{1/2} \qquad \text{for } t \geq c_3 r^2. \tag{2.33}$$

Since the constants in Lemma 2.13 are independent of n, the constants c_i in (2.33) are also independent of n. Since $M_\infty(t) \leq \liminf_{n\to\infty} M_n(t)$, (2.32) then follows. □

LEMMA 2.15. *Let $x \in G$, $r \geq 1$ and (x, r) be λ-good. Then*

$$c_1 t^{1/2} \leq E^x d(x, X_t) \leq c_2 t^{1/2} \qquad \textit{for } t \geq c_3 r^2. \tag{2.34}$$

PROOF. Since $d(x, X_t) \geq C_A^{-1}\widetilde{d}(x, X_t)$, the first inequality is clear. Let c_3 be as in Proposition 2.14, and $t = c_3 R_1^2$, so $R_1 \geq r$. Then if $A = B(x, R_1)^c$, using (2.20),

$$\begin{aligned} E^x d(x, X_t) &\leq \lambda R_1 + E^x\bigl(d(x, X_t); X_t \in A\bigr) \\ &\leq \lambda R_1 + E^x\bigl(\lambda\widetilde{d}(x, X_t); X_t \in A\bigr) \\ &\leq \lambda R_1 + \lambda c_4 t^{1/2} \leq c_5 t^{1/2}. \end{aligned}$$ □

The next few results follow quite closely along the lines of [1]. Let

$$\tau(x, r) = \inf\{t : d(x, X_t) \geq r\}.$$

LEMMA 2.16. *There exist constants c_1, c_2, c_3 such that if $R \geq c_1$ and*

$$(y, c_2 R) \textit{ is } \lambda\textit{-good for all } y \in B(x, R), \tag{2.35}$$

then if $t_0 = R^2/(2c_3)$

$$P^x\bigl(\tau(x, R/2) \leq t_0\bigr) \leq \tfrac{1}{2} \tag{2.36}$$

and hence for $t \geq 0$,

$$P^x\bigl(\tau(x, R) \leq t\bigr) \leq \frac{1}{2} + \frac{c_3 t}{R^2}. \tag{2.37}$$

PROOF. Write $\tau = \tau(x, R/2)$, and c_i' for the constants c_i in Lemma 2.15. Let $c_3 = 64(c_2')^2$. Choose c_2 so that $r = c_2 R$ satisfies $c_3' r^2 = t_0$, and c_1 so that $r \geq 1$. Then as (2.35) holds we can use Lemma 2.15 to bound $E^y d(y, X_s)$ for $s \geq t_0$, $y \in B(x, R)$. So,

$$\begin{aligned} c_2'(2t_0)^{1/2} &\geq E^x d(x, X_{2t_0}) \geq E^x\big(d(x, X_{t_0 \wedge \tau}) - d(X_{t_0\wedge\tau}, X_{2t_0})\big) \\ &\geq E^x 1_{(\tau \leq t_0)} d(x, X_\tau) - \sup_{y \in B(x,R)} \sup_{0 \leq s \leq t_0} E^y d(y, X_{2t_0 - s}) \\ &\geq P^x(\tau \leq t_0) R/2 - c_2'(2t_0)^{1/2}, \end{aligned}$$

and rearranging we obtain (2.36).

Inequality (2.37) now follows easily; if $t \leq t_0$, $P^x(\tau(x, R) \leq t) \leq P^x(\tau(x, R/2) \leq t_0) \leq \frac{1}{2}$ while if $t > t_0$, then the right-hand side of (2.37) is greater than 1. $\square$

To obtain the Gaussian upper bound on $q_t(x, y)$ we need to prove that the process X does not move too rapidly across a ball $B(x, R)$. We choose $r \ll R$, and use the fact that if X moves across $B(x, R)$ in the time interval $[0, t]$, then it has to move across many smaller balls of side r in the same period; the estimate (2.37) is enough to bound the probability of this. Our argument uses the following easily proved estimate:

LEMMA 2.17 (See [5], Lemma 1.1). *Let $\xi_1, \xi_2, \ldots, \xi_n$, V be nonnegative r.v. such that $V \geq \sum_1^n \xi_i$. Suppose that for some $p \in (0, 1)$, $a > 0$,*

$$P\big(\xi_i \leq t | \sigma(\xi_1, \ldots, \xi_{i-1})\big) \leq p + at, \qquad t > 0. \tag{2.38}$$

Then

$$\log P(V \leq t) \leq 2\left(\frac{ant}{p}\right)^{1/2} - n \log(1/p). \tag{2.39}$$

PROPOSITION 2.18. *There exists constants $c_1, \ldots, c_4$ such that if $x \in G$, $R \geq c_1$, $t \geq c_1 R$ and*

$$(z, c_2 t/R) \text{ is } \lambda\text{-good for all } z \in B(x, R), \tag{2.40}$$

then

$$P^x\big(\tau(x, R) < t\big) \leq c_3 e^{-c_4 R^2/t}. \tag{2.41}$$

PROOF. Let $1 \leq m < R/2$, and set $r = R/2m$. Define stopping times

$$S_0 = 0, \qquad S_i = \inf\{t \geq S_{i-1} : d(X_{S_{i-1}}, X_t) \geq r\}, \qquad i \geq 1.$$

Set $\tau_i = S_i - S_{i-1}$, and write $\mathcal{F}_t = \sigma(X_s, s \leq t)$ for the filtration of X. As $d(X_{S_i}, X_{S_{i+1}}) \leq r + 1 < 2r$, we have $S_m \leq \tau(x, R)$ and $X_{S_i} \in B(x, R)$ for $0 \leq i \leq m - 1$.

Suppose for the moment that m is such that we can apply Lemma 2.16 to control each τ_i. Then

$$P^x(\tau_i \le u | \mathcal{F}_{S_{i-1}}) \le \frac{1}{2} + \frac{c_5 u}{r^2}, \qquad u > 0, 1 \le i \le m, \tag{2.42}$$

so writing $p = \frac{1}{2}$, $a = c_5/r^2$ and using (2.39), we obtain

$$\begin{aligned} \log P^x\big(\tau(x,r) < t\big) &\le \log P^x\left(\sum_1^m \tau_i < t\right) \\ &\le 2(amt/p)^{1/2} - m \log p^{-1} \\ &\le -c_6 m\left(2 - \left(\frac{c_7 t m}{R^2}\right)^{1/2}\right) \\ &= -c_6 m\big(2 - (m/\kappa)^{1/2}\big), \end{aligned} \tag{2.43}$$

where $\kappa = R^2/(c_7 t)$. If κ is such that we can choose $m \in \mathbb{N}$ with $\kappa \le m < 2\kappa$, and so that (2.42) holds, then (2.43) implies (2.41).

We can choose c_1 so that $\kappa < R/2 - 1$. If $\kappa \le 1$ then, adjusting the constant c_3 appropriately, (2.41) is immediate. If $1 < \kappa < R/2 - 1$ then let $m = \lfloor \kappa \rfloor + 1 \le 2\kappa$. Then $\frac{1}{4} c_6(t/R) \le r \le \frac{1}{2} c_6(t/R)$, and so choosing c_2 suitably, (2.40) and Lemma 2.16 imply (2.42). $\square$

Recall that V_x is the smallest R such that (x, R) is λ-very good.

THEOREM 2.19. *Let $x, y \in G$, and write $D = d(x,y)$. Suppose that either $D \ge c_1 \vee V_x$ or $t \ge D^2$.*

(a) *If $c_2 D \le t$, then*

$$q_t(x,y) \le c_3 t^{-d/2} e^{-c_4 d(x,y)^2/t}. \tag{2.44}$$

(b) *If $c_2 D \ge t$, then*

$$q_t(x,y) \le c_3 \exp\big(-c_4 D\big(1 \vee \log(D/t)\big)\big). \tag{2.45}$$

PROOF. We need to consider various cases. First, if $t \ge D^2$, then (2.44) follows from (2.18). So we can suppose $D \ge V_x$. Note that by (2.20) $\widetilde{d}(x,y) \ge \lambda^{-1} D$.

If $t \le c_2 D$, then (2.45) follows from (2.10). If $c_2 D \le t \le c_6 D^2/\log D$, then Theorem 2.3 gives

$$q_t(x,y) \le c_8 \exp(-2c_7 D^2/t).$$

Choosing c_6 small enough we have $\exp(-c_7 D^2/t) \le t^{-d/2}$ and (2.44) follows.

It remains to consider the case $c_6 D^2/\log D \le t \le D^2$. Let $A_x = \{z : d(x,z) \le d(y,z)\}$, $A_y = G - A_x$, $t' = t/2$, $D' = D/2$. Note that $B(x, D') \subseteq A_x$. Then

$$\nu_x P^x(X_t = y) = \nu_x P^x(X_t = y, X_{t'} \in A_y) + \nu_x P^x(X_t = y, X_{t'} \in A_x). \tag{2.46}$$

To bound the first term in (2.46), and writing $\tau = \tau(x, D')$, we have

$$
\begin{aligned}
&P^x(X_t = y, X_{t'} \in A_y) \\
&\quad = P^x(\tau < t', X_{t'} \in A_y, X_t = y) \\
&\quad \le E^x\big(1_{\{\tau < t'\}} P^{X_\tau}(X_{t-\tau} = y)\big) \\
&\quad \le P^x\big(\tau(x, D') < t'\big) \sup_{z \in \partial B(x, D'), u \le t/2} q_{t-u}(z, y)\nu_y \\
&\quad \le c t^{-d/2} P^x\big(\tau(x, D') < t/2\big).
\end{aligned}
\tag{2.47}
$$

Similarly, using symmetry, for the second term in (2.46) we have

$$
\begin{aligned}
\nu_x P^x(X_t = y, X_{t'} \in A_x) &= \nu_y P^y(X_t = x, X_{t'} \in A_x) \\
&\le c t^{-d/2} P^y\big(\tau(y, D') < t/2\big).
\end{aligned}
\tag{2.48}
$$

It remains to verify that we can use Proposition 2.18 to bound the terms $P^z(\tau(z, D') < t/2)$ for $z = x, y$. Writing c_i' for the constants c_i in Proposition 2.18, taking c_1 large enough we have $D' \ge c_1'$, and $t' \ge c_1' D'$. As (x, V_x) is λ-very good, and $D \ge V_x$, we have that (z, r) is λ-good for $z \in B(x, 2D)$, and $r \ge (2D)^\eta$. We have $c_2' t'/D' \ge (2D)^\eta$ provided $t \ge c_8 D^{1+\eta}$, and since $t \ge c_6 D^2/\log D$ this holds by adjusting the constant c_1. So

$$
P^z\big(\tau(w, D') < t/2\big) \le c \exp(-c' D^2/t) \qquad \text{for } z = x, y,
\tag{2.49}
$$

and combining the estimates (2.46)–(2.49) completes the proof. $\square$

REMARK 2.20. This theorem does not give any bound for $q_t(x, y)$ when $D < c_1 \vee V_x$ and $t < D^2$. In this case we still have the global upper bound (2.18). In addition the "long-range" bounds in Theorem 2.3, bound $q_t(x, y)$ in terms of $\widetilde{d}(x, y)$, but we do not have a bound in terms of D.

The final result of this section is that, under fairly mild additional conditions, functions which are harmonic for the discrete time process $X_n, n \in \mathbb{Z}_+$ are also harmonic for the continuous time process $X_t, t \in \mathbb{R}_+$. At the end of Section 5 we will use this remark to note that the corrector constructed using the discrete time process also gives us a corrector for the continuous time process X_t.

Let $\widehat{X}$ be the discrete time process given by $\widehat{X}_n = X_n, n \in \mathbb{Z}_+$. Write

$$
\widehat{\mathcal{L}} f(x) = \sum_y q_1(x, y)\nu_y\big(f(y) - f(x)\big).
\tag{2.50}
$$

We say h is $\widehat{\mathcal{L}}$ harmonic if the sum in (2.50) converges absolutely for all x, and $\widehat{\mathcal{L}} h(x) = 0$ for all x. This implies that $(h(\widehat{X}_n), n \in \mathbb{Z}_+)$ is a P^x-martingale for each $x \in G$.

For $x \in G$ let $\kappa_x = \mu_x/\nu_x$ be the jump rate out of x by X. Set

$$
A(K) = \{y \in G : \kappa_y \le K\}.
\tag{2.51}
$$

LEMMA 2.21. *Let Γ satisfy Assumption 2.6. In addition suppose that there exist (x_0, r_0) such that (x_0, r_0) is λ-good, and that there exist R_0, K such when $R \geq R_0$ then every self avoiding path γ from x_0 to $B(x_0, R)^c$ contains at least $R^{1/2}$ points in $A(K)$. Let $h : G \to \mathbb{R}$ be $\widehat{\mathcal{L}}$ harmonic, and satisfy the growth condition*

$$|h(x)| \leq C_1 + C_1 d(x_0, x)^p \tag{2.52}$$

for some $p \in [1, \infty)$. Then $\mathcal{L}h = 0$, so that h is harmonic for X.

PROOF. By Lemma 2.10(b) we have that (x, r) is 2λ-good if $r/2 = r_0 \vee (1 + C_A\lambda)d(x_0, x)$. So by Lemma 2.12 there exists C_2 (depending on C_1 and λ) so that if $s \in [0, 1]$,

$$E^x|h(X_s)| \leq cr^{d+p} \leq C_2\big(r_0^{p+d} + d(x_0, x)^p\big); \tag{2.53}$$

it follows that $E^x h(X_t)$ is well defined for any $t \geq 0$. Set for $s \in [0, \infty)$

$$h_s(x) = E^x h(X_s).$$

To prove the lemma, it is sufficient to prove that $h = h_s$ for every s; this implies that $h(X_t)$ is a continuous time martingale and hence that $\mathcal{L}h = 0$. We have

$$h_{s+1}(x) = E^x(E^{X_s} h(X_1)) = E^x(h(X_s)) = h_s(x),$$

so $s \to h_s$ has period 1. We extend h_s by periodicity to $s \in \mathbb{R}$. Since $E^x h_s(X_1) = E^x h(X_{1+s}) = h_s(x)$, each h_s is $\widehat{\mathcal{L}}$-harmonic. Let

$$k(x) = \sup_{0 \leq s_1 \leq s_2 \leq 1} |h_{s_2}(x) - h_{s_1}(x)|;$$

note that by (2.53) we have

$$k(x) \leq 2 \sup_{s \leq 1} E^x|h(X_s)| \leq 2C_2\big(r^{p+d} + d(x_0, x)^p\big). \tag{2.54}$$

Fix $x \in G$, and write $\kappa = \kappa_x$. Write $P_{xy} = \mu_{xy}/\mu_x$ for the jump probabilities of X. Then by conditioning on the time of the first jump of X, if it occurs in $[0, 1]$, we obtain

$$h_s(x) = e^{-\kappa}h_s(x) + \sum_y P_{xy} \int_0^1 \kappa e^{-\kappa u} h_{s-u}(y)\, du.$$

So

$$\begin{aligned} &h_s(x)(1 - e^{-\kappa}) \\ &\quad = \sum_y P_{xy}\Big(\kappa \int_0^1 e^{-\kappa} h_{s-u}(y)\, du + \int_0^1 \kappa(e^{-\kappa u} - e^{-\kappa})h_{s-u}(y)\, du\Big) \\ &\quad = \sum_y P_{xy}\Big(\kappa \int_0^1 e^{-\kappa} h_u(y)\, du + \int_0^1 \kappa(e^{-\kappa u} - e^{-\kappa})h_{s-u}(y)\, du\Big). \end{aligned} \tag{2.55}$$

Then (2.55) implies that

$$k(x)(1-e^{-\kappa_x}) \le \sum_y P_{xy}k(y)\bigl(1-(1+\kappa_x)e^{-\kappa_x}\bigr). \tag{2.56}$$

So if $k(x)>0$, then there exists $y\sim x$ with $k(y)>k(x)$. Further, if $\kappa_x \le K$, then there exists $\delta>0$ (depending only on C_M and K) such that

$$k(y) \ge (1+\delta)k(x) \qquad \text{for some } y\sim x. \tag{2.57}$$

Suppose now that there exists x_1 with $k(x_1)>0$. Then there exists a nonintersecting infinite path γ_1 starting at x_1 on which k is strictly increasing. Let γ_2 be a shortest path from x_0 to a closest point y on γ to x_0, and let D be the length of γ_2. Combining γ_2 and the infinite segment of γ_1 starting at y, we obtain a path $\gamma=(x_0,z_1,\dots)$ for which $k(z_n)>0$ for all $n>D$. Let $R>R_0$, and let w_R be the first point in $\gamma\cap B(x_0,R)^c\cap A(K)$. Then $R_1=d(x_0,w_R)\ge R$. So, using (2.54), (2.57) and the condition on $A(K)$,

$$2C_2(r_0^{p+d}+R_1^p) \ge k(w_R) \ge (1+\delta)^{R_1^{1/2}-D}k(x_1),$$

which is a contradiction if R is large enough. □

3. Lower bounds and Harnack inequalities. Unlike the papers [10, 12, 34] we will need to make explicit use of heat kernel lower bounds in our proof of the invariance principle Theorem 1.1 (see Lemma 5.9).

In this section we specialize to the case when Γ is the d-dimensional Euclidean lattice, and μ_e are bond conductances with $\mu_e \ge 1$. We continue to assume that Assumptions 2.1 and 2.6 hold. Note that balls and distance are with respect to the graph distance on $\mathbb{Z}^d$.

We can follow the arguments in Section 5 of [1] fairly closely. First, as $\mu_{xy}\ge 1$ when $x\sim y$, by comparison with the standard Dirichlet form $\mathcal{E}_0$ on $\mathbb{Z}^d$ we have a weighted Poincaré inequality as in [1], Theorem 4.8.

THEOREM 3.1. *Let $B=B(x_0,R)$, $\rho_B(y)=d(y,B^c)$ and $\varphi(x)=R^2(R\wedge \rho_B(y))^2$. Then if $f:B\to\mathbb{R}$,*

$$\inf_a \sum_{x\in B}\bigl(f(x)-a\bigr)^2\varphi(x)\nu_x \le CR^2 \sum_{x,y\in B}\bigl(f(x)-f(y)\bigr)^2\varphi(x)\wedge\varphi(y)\mu_{xy}. \tag{3.1}$$

Using this, and the method of Fabes and Stroock [21] we obtain a lower bound of the form $q_t(x,y)\ge ct^{-d/2}$ when x,y are close enough together.

PROPOSITION 3.2. *Let $x_0\in\mathbb{Z}^d$ and $R\ge c_1$. Then provided*

$$(z,c_2R) \text{ is } \lambda\text{-good for all } z\in B(x_0,R), \tag{3.2}$$

we have

$$q_t(x_1,x_2)\ge c_1t^{-d/2} \qquad \text{for } x_1,x_2\in B(x_0,R/2),\ \tfrac{1}{8}R^2\le t\le R^2. \tag{3.3}$$

PROOF. This can be proved using the argument in [1], Proposition 5.1, with only minor changes. Note that we need to show that $\mathbb{P}^{x_1}(X_t \notin B(x_0, 2R/3)) \le \frac{1}{2}$ when $t = \theta R^2$ and θ is sufficiently small (see (5.2) and (5.9) in [1]). [There is a missing minus sign in exponential in the last line of (5.2).] This is done using Lemma 2.16, and so to satisfy (2.40) we need (3.2). □

THEOREM 3.3. *Let $x, y \in \mathbb{Z}^d$, $t > 0$, and write $D = d(x, y)$. Suppose that*

$$t \ge c_1 \vee V_x^2 \vee D^{1+\eta}. \tag{3.4}$$

Then

$$q_t(x, y) \ge c_2 t^{-d/2} e^{-c_3 d(x,y)^2/t}. \tag{3.5}$$

PROOF. The proof as in [1], Lemma 5.2, Theorem 5.3, follows by a standard chaining argument. We just give the details of the conditions on V_x, D and t needed to make this argument work.

First, if $D^2 < t$ then the lower bound in (3.5) is just $t^{-d/2}$, so we can use Proposition 3.2. We set $R = ct^{1/2}$. Then $t \ge V_x^2$ implies $R \ge V_x$, so $B(x, R)$ is λ-very good, and so as $cR \ge R^\eta$, (3.2) holds.

If $D^2 \ge t$ then we set $R = 2D$, $r = ct/D$. We apply Proposition 3.2 in a chain of balls $B_i = B(z_i, r)$ linking x and y. (See [1], Lemma 5.2, or [21] for details of the calculations.) Since $D \ge t^{1/2} \ge V_x$, we have that (x, R) is λ-very good, and hence that (z, cr') is λ-good for all $r' \ge R^\eta$, $z \in B(x, R)$. As $r = ct/D \ge c' R^\eta$, (3.2) holds for all the balls B_i. □

REMARK 3.4. 1. Note that the lower bounds do not extend to the range when $t \simeq D$. The difficulty is that if $t \simeq D$, then we need Proposition 3.2 for a chain of balls of radius $O(1)$ connecting x and y. The hypothesis "very good" is not enough to ensure this.

However, the chaining argument does not need (3.2) for all points in $B(x_0, R)$, but just for a suitable chain connecting x and y. In [1] this fact was used to obtain full Gaussian lower bounds. It is likely that the same approach will work for the random conductance model, but we do not pursue this point, since the bounds in Theorem 3.3 are enough for most applications.

2. A well-known theorem (see [23, 33]) states that for Brownian motion on a manifold Gaussian bounds are equivalent to two conditions: volume doubling plus a family of Poincaré inequalities. This theorem was extended to graphs in [18]. Since we have volume doubling (for ν) and the Poincaré inequalities hold (since $\mu_e \ge 1$), one might therefore ask if Theorems 2.19 and 3.3 follow immediately from known results.

However, it is clear that some conditions on μ_e are needed before Theorem 2.19 holds—one has to prevent X from moving a long distance in a very short time. In fact, examination of the theorems in [18, 23, 33] shows that in each case there is a "hidden" additional assumption which prevents the process from moving too quickly. For example, [18] considers a discrete time nearest neighbour random walk.

For $B \subset \mathbb{Z}^d$ let $q_t^B(x, y)$ be the transition density for the processes X killed on exiting from B.

LEMMA 3.5. *Let* (x_0, R) *be* λ*-very good. Then*

$$q_t^{B(x_0,R)}(x, y) \geq c_1 t^{-d/2}, \qquad x, y \in B(x_0, 3R/4), c_2 R^2 \leq t \leq R^2. \tag{3.6}$$

PROOF. Using Theorem 2.19 and Proposition 3.2, this follows, as in [1], Lemma 5.8, by the argument in [21], Lemma 5.1. $\square$

We now give a parabolic Harnack inequality (PHI) for X. The statement requires a little extra notation. If $A \subset \mathbb{Z}^d$ we write $\partial A = \{y : y \sim x \text{ for some } x \in A\}$ for the exterior boundary of A, and $\overline{A} = A \cup \partial A$. We call a function $u(t, x)$ *caloric* in a space–time region $Q = A \times (0, T) \subset [0, \infty) \times \mathbb{Z}^d$ if u is defined on $\overline{Q} = \overline{A} \times [0, T]$ and

$$\frac{\partial}{\partial t} u(t, x) = \mathcal{L}_V u(t, x), \qquad (t, x) \in Q.$$

Write $Q(x, R, T) = B(x, R) \times (0, T]$, $Q_-(x, R, T) = B(x, \frac{1}{2}R) \times [\frac{1}{4}T, \frac{1}{2}T]$ and $Q_+(x, R, T) = B(x, \frac{1}{2}R) \times [\frac{3}{4}T, T]$.

DEFINITION 3.6. We say the parabolic Harnack inequality (PHI) holds with constant C_P for $Q = Q(x, R, T)$ if whenever $u = u(t, x)$ is nonnegative and caloric on Q, then

$$\sup_{(t,x) \in Q_-(x,R,T)} u(t, x) \leq C_P \inf_{(t,x) \in Q_+(x,R,T)} u(t, x). \tag{3.7}$$

THEOREM 3.7 (Parabolic Harnack inequality). *There exists a constant* C_P *such that if* (x, R) *is* λ*-very good. Then the PHI holds with constant* C_P *in* $Q(x, R, R^2)$.

PROOF. Using the heat kernel bounds in Theorems 2.19 and 3.5, and Lemma 3.5, this follows by the same argument as in [3], Theorem 3.1. $\square$

4. Heat kernel bounds for the RCM. In this section we prove Theorem 1.2. Let E_d be the edges of the Euclidean lattice $\mathbb{Z}^d$, and let $\Omega = [1, \infty]^{E_d}$. Let $\mathbb{P}$ be a probability measure on Ω which makes the coordinates i.i.d. with a law on $[1, \infty)$. We set $\mu_e(\omega) = \omega(e)$ for $e \in E_d$, and for each $\omega \in \Omega$ we consider the random walk X on the graph $(\mathbb{Z}^d, E_d)$ with conductances $\mu_e(\omega)$.

Using the notation of Section 2 we take $\nu_x = 1$ for all x, so that we can take $C_M = 1$. We write P^x_ω for the law of X started at x, and

$$q^\omega_t(x, y) = P^x_\omega(X_t = y)$$

for the transition density of X.

LEMMA 4.1. *The graph* $(\mathbb{Z}^d, E_d)$, *conductances* μ_e *and random walk* X *satisfy Assumptions* 2.1 *and* 2.6 *with* $\mathbb{P}$*-probability* 1.

PROOF. Assumption 2.1 is immediate; note we can take $C_D = 2d$. Setting $C_V = 2^d$ Assumption 2.6(1) is also immediate.

Since we have $\mu(e) \geq 1$ for all edges in $\mathbb{Z}^d$, if $\mathcal{E}_0$ is the Dirichlet form of the standard continuous time random walk on $\mathbb{Z}^d$, then $\mathcal{E}(f, f) \geq \mathcal{E}_0(f, f)$, so that the standard Nash inequality on $\mathbb{Z}^d$ (see [14]) implies Assumption 2.6(2) with a constant C_N depending only on d. $\square$

In what follows we set $C_A = 1$; thus none of the constants C_D, C_A, C_M, C_N, C_V depend on the law of μ_e (apart from the fact that $\mathbb{P}(\mu_e \in [1, \infty)) = 1$).

Let $d(x, y)$ be the graph metric on $(\mathbb{Z}^d, E_d)$, and $\widetilde{d} = \widetilde{d}(\omega)$ be the metric given by (2.6); as in the previous sections we write $\widetilde{B}(x, r)$ for balls in the $\widetilde{d}$ metric. Write $B_E(x, r) = \{y \in \mathbb{R}^d : |x - y| < r\}$ for the Euclidean ball center x and radius r.

LEMMA 4.2. *There exists a constant* $\lambda_0 > 0$ *such that*

$$\mathbb{P}\big(\widetilde{B}(0, r) \subset B_E(0, \lambda_0 r)\big) \geq 1 - c_1 e^{-c_2 r}. \tag{4.1}$$

PROOF. We use results on first passage percolation from [25]. As in [25] let $b_{0,n}$ be the first time $\widetilde{B}(0, t)$ reaches the hyperplane $\{x_1 = n\}$. Using [25], Theorem 2.18, there exists μ_0 such that $\lim_n n^{-1} b_{0,n} = \mu_0$, a.s. and in L^1. By [25], Theorem 1.15, we have $\mu_0 > 0$. By [25], Theorem 5.2, there exist $c_3, c_4 > 0$ such that

$$\mathbb{P}\big(b_{0,n} < \tfrac{1}{2} n \mu_0\big) \leq c_3 e^{-c_4 n}, \qquad n \geq 0. \tag{4.2}$$

The times for $\widetilde{B}(0, t)$ to hit each hyperplane $\{x_i = \pm n\}$, for $i = 1, \ldots, d$ have the same law as $b_{0,n}$, so we deduce

$$\mathbb{P}\big(\widetilde{B}\big(0, \tfrac{1}{2}\mu_0 n\big) \subset [-n, n]^d\big) \geq 1 - 2d c_3 e^{-c_4 n}, \qquad n \geq 0, \tag{4.3}$$

and (4.1) follows easily. $\square$

Note that λ_0 does depend on the law of μ_e. We fix $\eta \in (0, 1)$, and define *good* and *very good* as in Section 2, with λ replaced by λ_0; and we write V_x for the smallest integer such that (x, V_x) is very good.

THEOREM 4.3. (a)

$$\mathbb{P}\big((x,r) \text{ is not good}\big) \le ce^{-cr}, \qquad r \ge r_0.$$

(b)

$$\mathbb{P}(V_x \ge n) \le c\exp(-cn^\eta). \tag{4.4}$$

PROOF. Let $G(y,r) = \{(y,r) \text{ is good}\}$, and $F(R) = \{(y,r) \text{ is good for all } y \in B(0,2R), r \ge R^\eta\}$. Then

$$\mathbb{P}(G(y,r)^c) \le \sum_{n=r}^{\infty} ce^{-cn} \le ce^{-cr}.$$

So,

$$\mathbb{P}(F(R)^c) \le cR^d \sum_{k=R^\eta}^{\infty} c_3 e^{-c_2 k} \le c\exp(-cR^\eta),$$

and since $\{V_0 \ge n\} = \bigcup_n^\infty F(k)^c$, (b) follows. □

Using Lemma 2.11 we obtain the following:

COROLLARY 4.4. *X is conservative with* $\mathbb{P}$*-probability* 1.

COROLLARY 4.5. *Let* $x \in \mathbb{Z}^d$. *Then*

$$\lim_{M\to\infty} \lim_{t\to\infty} P_\omega^0\big(|X_t| \ge M\sqrt{t}\big) = 0, \qquad \mathbb{P}\text{-a.s.} \tag{4.5}$$

PROOF. By Lemma 2.15, for $t \ge cV_0^2(\omega)$,

$$P_\omega^0\big(|X_t| \ge M\sqrt{t}\big) \le cM^{-1}t^{-1/2}E_\omega^0 d(0,X_t) \le cM^{-1},$$

and (4.5) follows. □

THEOREM 4.6. *There exist r.v.* U_x, $x \in \mathbb{Z}^d$, *such that*

$$\mathbb{P}\big(U_x(\omega) \ge n\big) \le c_1 \exp(-c_2 n^\eta), \tag{4.6}$$

and if $|x-y| \vee t^{1/2} \ge U_x$, *then*

$$q_t^\omega(x,y) \le c_3 t^{-d/2} e^{-c_4|x-y|^2/t} \qquad \text{when } t \ge |x-y|, \tag{4.7}$$

$$q_t^\omega(x,y) \le c_3 \exp\big(-c_4|x-y|\big(1 \vee \log(|x-y|/t)\big)\big) \qquad \text{when } t \le |x-y|. \tag{4.8}$$

Further,

$$q_t^\omega(x,y) \ge c_6 t^{-d/2} e^{-c_7|x-y|^2/t} \qquad \text{if } t \ge U_x^2 \vee |x-y|^{1+\eta}. \tag{4.9}$$

PROOF. We take $U_x = c_8(V_x + 1)$ where $c_8 \geq 1$. The bounds then follow from Theorems 2.19 and 3.3. [Note that the bounds (4.7) and (4.8) are of the same form if $d(x, y) \leq t \leq cd(x, y)$.] We use the constant c_8 to adjust between the the Euclidean metric $|x - y|$ and the graph metric $d(x, y)$, and to absorb the conditions $d(x, y) \geq c$ and $t \geq c$ into (4.6). $\square$

THEOREM 4.7. *There exists a constant C_P and r.v. $U_x, x \in \mathbb{Z}^d$ with*

$$\mathbb{P}\big(U_x(\omega) \geq n\big) \leq c_1 \exp(-c_2 n^\eta), \tag{4.10}$$

such that if $R \geq U_x$ then a PHI with constant C_P holds for $Q(x, R, R^2)$.

PROOF. This is immediate from Theorem 3.7 and (4.4). $\square$

The PHI implies an elliptic Harnack inequality (EHI), which holds for the CSRW as well as the VSRW. A function h is harmonic on $A \subset \mathbb{Z}^d$ if it is defined on $\overline{A}$ and $\mathcal{L}_V h(x) = 0$ [or equivalently $\mathcal{L}_C h(x) = 0$] for $x \in A$.

COROLLARY 4.8. *There exists a constant C_H and r.v. $U_x, x \in \mathbb{Z}^d$ with*

$$\mathbb{P}\big(U_x(\omega) \geq n\big) \leq c_1 \exp(-c_2 n^\eta), \tag{4.11}$$

such that if $R \geq U_x$, then an EHI with constant C_H holds for $B(x, R)$; if $h \geq 0$ is harmonic in $B(x, R)$, then

$$\sup_{y \in B(x,R/2)} h(y) \leq C_H \inf_{y \in B(x,R/2)} h(y). \tag{4.12}$$

We have the following averaged bounds:

THEOREM 4.9. (a) *Let $x, y \in \mathbb{Z}^d$ and $t \geq c_1 \vee |x - y|^{1+\eta}$. Then*

$$c_2 t^{-d/2} e^{-c_3|x-y|^2/t} \leq \mathbb{E} q_t^\omega(x, y) \leq c_4 t^{-d/2} e^{-c_5|x-y|^2/t}. \tag{4.13}$$

(b) *We have*

$$\mathbb{E} E_\omega^0 |X_t|^2 \leq c_6 t, \qquad t \geq 1. \tag{4.14}$$

PROOF. (a) Let $D = |x - y|$. Choose c_1 so that $\mathbb{P}(U_x > c_1^{1/2}) < \frac{1}{2}$. Then if $t \geq c_1 \vee D^{1+\eta}$, by (4.9),

$$\mathbb{E} q_t^\omega(x, y) \geq \mathbb{E}\big(q_t^\omega(x, y); U_x^2 < c_1\big) \geq \tfrac{1}{2} c t^{-d/2} e^{-cD^2/t}.$$

For the upper bound, let $\eta' = 1 - \eta$, and R_x' be the r.v. given in Theorem 4.6 using η' instead of η. Then by (4.7) and (4.6),

$$\begin{aligned}\mathbb{E} q_t^\omega(x, y) &= \mathbb{E}\big(q_t^\omega(x, y); R_x' > D\big) + \mathbb{E}\big(q_t^\omega(x, y); R_x' \leq D\big) \\ &\leq c t^{-d/2} e^{-cD^{\eta'}} + c t^{-d/2} e^{-cD^2/t}.\end{aligned}$$

Since the second term is larger, the upper bound in (4.13) follows.

(b) We have

$$E^0_\omega |X_t|^2 = \sum_x |x|^2 q^\omega_t(0,x). \tag{4.15}$$

We split the sum in (4.15) into three parts. First,

$$\sum_{|x|<U_0} |x|^2 q^\omega_t(0,x) \le U_0^2. \tag{4.16}$$

Next, using (4.7),

$$\sum_{U_0<|x|<ct} |x|^2 q^\omega_t(0,x) \le \sum_{U_0<|x|<ct} |x|^2 c t^{-d/2} e^{-c|x|^2/t} \le ct. \tag{4.17}$$

Finally, using (4.8),

$$\sum_{ct\vee U_0\le|x|} |x|^2 q^\omega_t(0,x) \le \sum_{ct\le|x|} |x|^2 c e^{-c|x|} \le c'. \tag{4.18}$$

Combining (4.15), (4.16) and (4.17) gives

$$E^0_\omega |X_t|^2 \le ct + c' U_0^2,$$

and as by (4.6) $\mathbb{E}U_0^2 < \infty$ and $t \ge 1$, we obtain (4.14). $\square$

REMARK 4.10. Combining Lemma 2.8, Theorems 4.6 and 4.9 completes the proof of Theorem 1.2.

Now let

$$X^{(\varepsilon)}_t = \varepsilon X_{t/\varepsilon^2}, \qquad 0 < \varepsilon \le 1. \tag{4.19}$$

THEOREM 4.11. *Let $T > 0$, $\delta > 0$, $r > 0$. Then*

$$\lim_{R\to\infty} \sup_\varepsilon P^0_\omega\Big(\sup_{s\le T} |X^{(\varepsilon)}_s| > R\Big) \to 0, \tag{4.20}$$

$$\lim_{\delta\to 0} \limsup_{\varepsilon\to 0} P^0_\omega\Big(\sup_{|s_1-s_2|\le\delta, s_i\le T} |X^{(\varepsilon)}_{s_2} - X^{(\varepsilon)}_{s_1}| > r\Big) = 0. \tag{4.21}$$

PROOF. By Theorem 4.6, if $R \ge U_0$, then

$$P^0_\omega\Big(\sup_{s\le T} X_s \ge R\Big) \le c\exp(-cR^2/T).$$

So if $R \ge U_0$, then $R/\varepsilon \ge U_0$ and

$$P^0_\omega\Big(\sup_{s\le T} |X^{(\varepsilon)}_s| > R\Big) = P^0_\omega\Big(\sup_{s\le T/\varepsilon^2} |X_s| \ge R/\varepsilon\Big) \le c\exp(-cR^2/T),$$

proving (4.20).

To prove (4.21) write

$$p(T,\delta,r) = P_\omega^0\Big(\sup_{|s_1-s_2|\le\delta, s_i\le T} |X_{s_2}-X_{s_1}|>r\Big), \tag{4.22}$$

so that

$$P_\omega^0\Big(\sup_{|s_1-s_2|\le\delta, s_i\le T} \big|X^{(\varepsilon)}_{s_2}-X^{(\varepsilon)}_{s_1}\big|>r\Big) = p(T/\varepsilon^2, \delta/\varepsilon^2, r/\varepsilon).$$

We begin by bounding $p(T,\delta,r)$ for fixed T, δ and r. Let $\kappa \in (0,\frac{1}{2})$, $U_R^* = \max_{x\in B(0,R)} U_x$, and $H(R)=\{U_R^*\le R^\kappa\}$. Then

$$\mathbb{P}(H(R)^c) \le cR^d \exp(-cR^{\kappa\eta}), \tag{4.23}$$

so by Borel–Cantelli there exists $R_0=R_0(\omega)$ such that $\omega\in H(R)$ for all $R\ge R_0$.

Let

$$Z_k = \sup_{0\le s\le\delta} |X_{k\delta+s}-X_{k\delta}|. \tag{4.24}$$

Then if $K=\lfloor T/\delta\rfloor$ and $Z^*=\max_{0\le k\le K} Z_k$, it is enough to control Z^* since

$$\sup_{|s_1-s_2|\le\delta, s_i\le T} |X_{s_2}-X_{s_1}| \le 2Z^*.$$

Let $R\ge 1$. Then

$$P_\omega^0(Z^*\ge r) \le P_\omega^0\big(\tau(0,R)\le T\big) + P_\omega^0\big(Z^*\ge r, \tau(0,R)>T\big). \tag{4.25}$$

By Proposition 2.18 we have

$$P_\omega^0\big(\tau(0,R)\le T\big) \le c\exp(-cR^2/T), \tag{4.26}$$

provided that $(0,R)$ is very good. For this it is sufficient that $R\ge U_0(\omega)$. Now,

$$\begin{aligned} P_\omega^0\big(Z^*\ge r, \tau(0,R)>T\big) &\le \sum_{k=0}^K P_\omega^0\big(Z_k\ge r, X_{k\delta}\in B(0,R)\big) \\ &\le \sum_{k=0}^K \sum_{y\in B(0,R)} P_\omega^y\big(\tau(y,r)<\delta\big) P_\omega^0(X_{k\delta}=y). \end{aligned}$$

Again by Proposition 2.18, for $y\in B(0,R)$,

$$P_\omega^y\big(\tau(y,r)<\delta\big) \le c\exp(-cr^2/\delta), \tag{4.27}$$

provided $r\ge U_R^*$. This will hold if $R\ge R_0(\omega)$ and $r\ge R^\kappa$. Combining (4.25), (4.26), (4.27), we obtain

$$P_\omega^0(Z^*\ge r) \le c\exp(-cR^2/T) + c(T/\delta)\exp(-cr^2/\delta), \tag{4.28}$$

provided $R\ge R_0(\omega)$ and $r\ge R^\kappa$.

Hence

$$p(T/\varepsilon^2, \delta/\varepsilon^2, 2r/\varepsilon) \le c\exp(-cR^2/T) + c(T/\delta)\exp(-cr^2/\delta), \tag{4.29}$$

provided $R > \varepsilon R_0$ and $r \ge R^\kappa \varepsilon^{1-\kappa}$. For fixed r, δ choose R so that $R \ge R_0$ and $R^2/T \ge r^2/\delta$. Then

$$p(T/\varepsilon^2, \delta/\varepsilon, 2r/\varepsilon) \le cT\delta^{-1}\exp(-cr^2/\delta) \qquad \text{when } \varepsilon^{1-\kappa} \le rR^{-\kappa}.$$

Hence

$$\limsup_{\varepsilon\to 0} p(T/\varepsilon^2, \delta/\varepsilon, 2r/\varepsilon) \le cT\delta^{-1}\exp(-cr^2/\delta),$$

and (4.21) follows. □

For $n \in \mathbb{N}$ let $\widehat{X}_n = X_n$, and set

$$\widehat{X}_t^{(\varepsilon)} = \varepsilon \widehat{X}_{\lfloor t/\varepsilon^2 \rfloor}, \qquad 0 < \varepsilon \le 1. \tag{4.30}$$

LEMMA 4.12. *For any $u > 0$,*

$$\lim_{\varepsilon\to 0} P_\omega^0\Big(\sup_{0\le s\le T} |\widehat{X}_s^{(\varepsilon)} - X_s^{(\varepsilon)}| > u\Big) = 0. \tag{4.31}$$

PROOF. In the notation of the previous theorem, it is sufficient to bound $p(T/\varepsilon^2, 1, u/\varepsilon)$; using (4.29) we have

$$p(T/\varepsilon^2, 1, u/\varepsilon) \le cT\varepsilon^{-2}\exp(-cu^2/\varepsilon^2),$$

provided there exists an R with $R \ge \varepsilon R_0(\omega)$, $R^2 \ge Tu^2\varepsilon^{-2}$ and $u/\varepsilon \ge R^\kappa$. Setting $R = T^{1/2}u/\varepsilon$, we need $uT^{1/2} \ge \varepsilon^2 R_0(\omega)$, and $u^{1-\kappa} \ge \varepsilon^{1-\kappa}T^{\kappa/2}$, so these bounds hold for all sufficiently small ε. □

5. Invariance principle. In this section we prove the invariance principle Theorem 1.1. We assume that the conductances μ_e are defined on the space $(\Omega, \mathbb{P})$ where

$$\Omega = [1, \infty]^{E_d}.$$

We write $\mu_e(\omega) = \omega(e)$ for the coordinate maps, and make the following assumptions on the environment (μ_e).

ASSUMPTION 5.1. (1) (μ_e) is stationary, ergodic, and invariant under symmetries of $\mathbb{Z}^d$.

(2) $\mu_e \in [1, \infty)$ for all $e \in E_d$, $\mathbb{P}$-a.s.

(3) The conclusions of Theorem 1.2 hold for the VSRW associated with (μ_e).

As explained in the Introduction, our basic approach is to construct the "corrector" $\chi : \Omega \times \mathbb{Z}^d \to \mathbb{R}^d$ so that, for $\mathbb{P}$-a.a. ω the discrete time process

$$\widehat{M}_n = \widehat{X}_n - \chi(\omega, \widehat{X}_n) \tag{5.1}$$

is a P^0_ω-martingale with respect to the filtration $\widehat{\mathcal{F}}_n = \sigma(\widehat{X}_k, 0 \le k \le n)$.

The key steps in the proof of the invariance principle are:

1. Tightness (a consequence of Theorem 3.5);
2. The invariance principle for the martingale part. This is standard and follows from the ergodicity of our environment (see [29], proof of Theorem 2.6);
3. The almost sure control of the corrector, for which we use the the ergodicity of the environment, the properties (4.11) and (4.12) and the quenched heat kernel estimates in Theorem 1.2 (see [34], or [12], Theorem 2.3). Note that all we need here is the ergodicity of the environment; ergodicity under the action of each direction as stated in [34], Remark 1.3, is not required since one can use the cocycle property of the corrector (see [13, 26]).

We now give the details. Let

$$\Omega_0 = \{\omega : \omega(e) \in [1, \infty) \text{ for all } e\}.$$

Since $\omega(e)$ satisfies Assumption 5.1 we have $\mathbb{P}(\Omega_0) = 1$. We write $\omega = (\omega(e), e \in E_d)$, and $\omega(x, y) = \omega(\{x, y\})$. For $x \in \mathbb{Z}^d$ define $T_x : \Omega \to \Omega$ by

$$T_x(\omega)(z, w) = \omega(z + x, w + x).$$

Let X be the VSRW with generator $\mathcal{L}_V$ given by (1.3), and $q_t^\omega(x, y)$ be the transition density of X. As $\nu_x \equiv 1$ is the invariant measure for X,

$$q_t^\omega(x, y) = P^x_\omega(X_t = y) = q_t^\omega(y, x).$$

Write

$$Q_{xy}(\omega) = q_1^\omega(x, y), \qquad Q^{(n)}_{xy}(\omega) = q_n^\omega(x, y), \qquad x, y \in \mathbb{Z}^d; \tag{5.2}$$

and note that $Q_{xy} \le 1$ for all x, y, with $\sum_y Q_{xy} = 1$. We have

$$Q^{(n)}_{xy} \circ T_z = Q^{(n)}_{x+z,y+z}, \qquad Q^{(n)}_{xy} = Q^{(n)}_{yx}. \tag{5.3}$$

We define the process Z, which gives the "environment seen from the particle," by

$$Z_t = T_{X_t}\omega, \qquad t \in [0, \infty), \tag{5.4}$$

and define the discrete time process $\widehat{Z}$ by $\widehat{Z}_n = Z_n$, $n \in \mathbb{Z}_+$.

Let $L^p = L^p(\Omega, \mathbb{P})$. For $F \in L^2$ write $F_x = F \circ T_x$. Then $\widehat{Z}$ has generator

$$\widehat{L}F(\omega) = \sum_{x \in \mathbb{Z}^d} Q_{0x}(\omega)\big(F_x(\omega) - F(\omega)\big).$$

Set

$$\widehat{\mathcal{E}}(F, G) = \mathbb{E} \sum_{x \in \mathbb{Z}^d} Q_{0x}(F - F_x)(G - G_x).$$

LEMMA 5.2. *We have for $F \in L^1$,*

$$\mathbb{E}F = \mathbb{E}F_x, \tag{5.5}$$

$$\mathbb{E}(Q_{0x}F_x) = \mathbb{E}(Q_{0,-x}F). \tag{5.6}$$

PROOF. Since $\mathbb{P}$ is invariant by T_x the first relation is immediate. As $(Q_{0,x})_{-x} = Q_{-x,0} = Q_{0,-x}$ by (5.3), $\mathbb{E}(Q_{0x}F_x) = \mathbb{E}((Q_{0x})_{-x}F) = \mathbb{E}(Q_{0,-x}F)$, proving (5.5). $\square$

LEMMA 5.3. *If $F, G \in L^2$, then $\widehat{\mathcal{E}}(F, F) < \infty$, $\widehat{\mathcal{E}}(F, G)$ is defined, and $\widehat{L}F \in L^2$.*

PROOF. Let $F \in L^2$. Then

$$\begin{aligned}
\widehat{\mathcal{E}}(F, F) &= \mathbb{E}\sum_{x\in\mathbb{Z}^d} Q_{0x}(F - F_x)^2 \\
&\le 2\mathbb{E}\sum_{x\in\mathbb{Z}^d} Q_{0x}(F^2 + F_x^2) \\
&= 2\mathbb{E}F^2 + 2\mathbb{E}\sum_{x\in\mathbb{Z}^d} Q_{0x}F_x^2 \\
&= 2\mathbb{E}F^2 + 2\mathbb{E}\sum_{x\in\mathbb{Z}^d} Q_{0,-x}F^2 = 4\|F\|_2^2.
\end{aligned}$$

Hence $\widehat{\mathcal{E}}(F, G)$ is defined for $F, G \in L^2$. Also, if $F \in L^2$,

$$\begin{aligned}
\mathbb{E}|\widehat{L}F|^2 &= \mathbb{E}\sum_{x,y} Q_{0x}Q_{0y}(F_x - F)(F_y - F) \\
&\le \mathbb{E}\biggl[\biggl(\sum_{x,y} Q_{0x}Q_{0y}(F_x - F)^2\biggr)^{1/2}\biggl(\sum_{x,y} Q_{0x}Q_{0y}(F_y - F)^2\biggr)^{1/2}\biggr] \\
&= \widehat{\mathcal{E}}(F, F) \le 4\|F\|_2^2.
\end{aligned}$$

$\square$

LEMMA 5.4. *Let $F, G \in L^2$. Then*

$$\mathbb{E}(G\widehat{L}F) = -\widehat{\mathcal{E}}(F, G). \tag{5.7}$$

PROOF. Using (5.5) we have

$$\mathbb{E}\bigl(Q_{0,-x}G(F_{-x} - F)\bigr) = \mathbb{E}\bigl(Q_{0x}G_x(F - F_x)\bigr). \tag{5.8}$$

So

$$\begin{aligned}\mathbb{E}(G\widehat{L}F) &= \sum_{x\in\mathbb{Z}^d} \mathbb{E}GQ_{0x}(F_x - F)\\ &= \frac{1}{2}\sum_{x\in\mathbb{Z}^d} \mathbb{E}GQ_{0x}(F_x - F) + \frac{1}{2}\sum_{x\in\mathbb{Z}^d} \mathbb{E}GQ_{0,-x}(F_{-x} - F)\\ &= \frac{1}{2}\sum_{x\in\mathbb{Z}^d} \mathbb{E}Q_{0x}(GF_x - GF + G_xF - G_xF_x) = -\widehat{\mathcal{E}}(F,G),\end{aligned}$$

where we used (5.8) in the last line. □

Now we look at "vector fields." We define for $G = G(\omega, x) : \Omega \times \mathbb{Z}^d \to \mathbb{R}$,

$$\overline{\mathbb{E}}G = \sum_x \mathbb{E}Q_{0x}G(\cdot, x).$$

DEFINITION. We say $G(\omega, x)$ has the *cocycle property* (see [13, 26]) if

$$G(T_x\omega, y - x) = G(\omega, y) - G(\omega, x), \qquad \mathbb{P}\text{-a.s.} \tag{5.9}$$

Let $\mathcal{H} = \overline{L}^2$ be the set of vector fields G with the cocycle property and $\|G\|^2 = \overline{\mathbb{E}}G^2 < \infty$.

LEMMA 5.5. *Let $G = G(\omega, x) \in \overline{L}^2$.*
(a) $G(\omega, 0) = 0$, *and* $G(T_x\omega, -x) = -G(\omega, x)$.
(b) *If $x_0, x_1, \dots, x_n \in \mathbb{Z}^d$, then*

$$\sum_{i=1}^{n} G(T_{x_{i-1}}\omega, x_i - x_{i-1}) = G(\omega, x_n) - G(\omega, x_0). \tag{5.10}$$

PROOF. (a) follows immediately from the definition. For (b), as G has the cocycle property

$$G(T_{x_{i-1}}\omega, x_i - x_{i-1}) = G(\omega, x_i) - G(\omega, x_{i-1}),$$

giving (5.10). □

It is easy to check the following:

LEMMA 5.6. *$\overline{L}^2$ is a Hilbert space.*

For $F \in L^2$ we set

$$\nabla F(\omega, x) = F(T_x\omega) - F(\omega).$$

LEMMA 5.7. *If $F \in L^2$, then $\nabla F \in \overline{L}^2$.*

PROOF. First,

$$\overline{\mathbb{E}}|\nabla F|^2 = \sum_x \mathbb{E} Q_{0x}(F_x - F)^2 = \widehat{\mathcal{E}}(F, F) < \infty.$$

Also,

$$\begin{aligned}\nabla F(T_x\omega, y - x) &= F(T_{y-x}T_x\omega) - F(T_x\omega)\\ &= F(T_y\omega) - F(T_x\omega) = \nabla F(\omega, y) - \nabla F(\omega, x),\end{aligned}$$

so ∇F has the cocycle property. □

LEMMA 5.8. *Let $G \in \overline{L}^2$. Then*

$$\mathbb{E} \sum_x Q_{0x}^{(n)} G(\omega, x)^2 \le n\|G\|_2^2. \tag{5.11}$$

PROOF. Write a_n^2 for the left-hand side of (5.11). Then using (5.9),

$$a_n^2 = \mathbb{E} \sum_x \sum_y Q_{0x}^{(n-1)} Q_{xy}\bigl(G(T_x\omega, y - x) + G(\omega, x)\bigr)^2. \tag{5.12}$$

We now expand the final square in (5.12) and compute the three terms separately. We have

$$\begin{aligned}&\mathbb{E} \sum_x \sum_y Q_{0x}^{(n-1)}(\omega) Q_{xy}(\omega) G(T_x\omega, y - x)^2\\ &\quad = \mathbb{E} \sum_x \sum_y Q_{-x,0}^{(n-1)}(T_x\omega) Q_{0,y-x}(T_x\omega) G(T_x\omega, y - x)^2\\ &\quad = \mathbb{E} \sum_x \sum_z Q_{-x,0}^{(n-1)}(\omega) Q_{0,z}(\omega) G(\omega, z)^2\\ &\quad = \mathbb{E} \sum_z Q_{0,z}(\omega) G(\omega, z)^2 = \|G\|^2.\end{aligned}$$

Also,

$$\mathbb{E} \sum_x \sum_y Q_{0x}^{(n-1)}(\omega) Q_{xy}(\omega) G(\omega, x)^2 = \mathbb{E} \sum_x Q_{0x}^{(n-1)}(\omega) G(\omega, x)^2 = a_{n-1}^2.$$

Finally,

$$\begin{aligned}&\mathbb{E} \sum_x \sum_y Q_{0x}^{(n-1)}(\omega) Q_{xy}(\omega) G(\omega, x) G(T_x\omega, y - x)\\ &\quad = \mathbb{E} \sum_x \sum_z Q_{0x}^{(n-1)}(\omega) Q_{0,z}(T_x\omega) G(\omega, x) G(T_x\omega, z)\\ &\quad \le \Bigl(\mathbb{E} \sum_x \sum_z Q_{0x}^{(n-1)}(\omega) Q_{0,z}(T_x\omega) G(\omega, x)^2\Bigr)^{1/2}\end{aligned}$$

$$\times \left(\mathbb{E}\sum_x \sum_z Q_{0x}^{(n-1)}(\omega)Q_{0,z}(T_x\omega)G(T_x\omega, z)^2\right)^{1/2}$$
$$= a_{n-1}\|G\|.$$

Thus $a_n \le a_{n-1} + \|G\|$, and so $a_n \le n\|G\|$. □

Note that the following lemma uses the heat kernel lower bounds.

LEMMA 5.9. *Let $G \in \overline{L}^2$ and $1 \le p < 2$. Then there exists a constant $c_p < \infty$ such that*

$$(\mathbb{E}|G(\cdot, x)|^p)^{1/p} \le (c_p|x|)\|G\|. \tag{5.13}$$

It follows that, $\mathbb{P}$-*a.s.*,

$$\lim_{n\to\infty} \max_{|x|\le n} \frac{|G(\omega, x)|}{n^{d+4}} = 0. \tag{5.14}$$

PROOF. By (5.9) and the triangle inequality we have

$$(\mathbb{E}|G(\cdot, x)|^p)^{1/p} \le |x|(\mathbb{E}|G(\cdot, e_1)|^p)^{1/p};$$

so it is enough to bound $\mathbb{E}|G(\cdot, e_1)|^p$. By Theorem 1.2 there exists an integer valued random variable W_0 with $W_0 \ge 1$ such that $\mathbb{P}(W_0 = n) \le c_1 \exp(-c_2 n^\delta)$ for some $\delta > 0$ and $q_t^\omega(0, x) \ge c_3 t^{-d/2}$ for $t \ge W_0$. Write $\xi_n = q_n^\omega(0, e_1)$. Then

$$\mathbb{E}|G(\cdot, e_1)|^p = \sum_{n=1}^\infty \mathbb{E}|G(\cdot, e_1)|^p 1_{(W_0=n)}. \tag{5.15}$$

Let $\alpha = 2/p$, and let $\alpha' = 2/(2 - p)$ be its conjugate index. Then using Hölder's inequality and (5.11),

$$\begin{aligned}
&\mathbb{E}|G(\cdot, e_1)|^p 1_{(W_0=n)}\\
&\quad = \mathbb{E}\big(\xi_n^{1/\alpha}|G(\cdot, e_1)|^p \xi_n^{-1/\alpha} 1_{(W_0=n)}\big)\\
&\quad \le (\mathbb{E}\xi_n G(\cdot, e_1)^2)^{1/\alpha}\big(\mathbb{E}\xi_n^{-\alpha'/\alpha} 1_{(W_0=n)}\big)^{1/\alpha'}\\
&\quad \le \left(\mathbb{E}\sum_y Q_{0y}^{(n)} G(0, y)^2\right)^{1/\alpha} \big((c_3 n^{-d/2})^{-\alpha'/\alpha} c_1 \exp(-c_2 n^\delta)\big)^{1/\alpha'}\\
&\quad \le (n\|G\|^2)^{1/\alpha} c_4 n^{d/2\alpha} \exp(-c_5 n^\delta)\\
&\quad = c_4 n^{(d+2)/2\alpha} \exp(-c_5 n^\delta)\|G\|^p.
\end{aligned}$$

Summing the series in n we obtain (5.13).

Using (5.13) with $p=1$ we have

$$\mathbb{P}\Big(\max_{|x|\le n}|G(\omega,x)|>\lambda_n\Big)\le(2n)^d\max_{|x|\le n}\mathbb{P}\big(|G(\omega,x)|>\lambda_n\big)$$
$$\le cn^d\lambda_n^{-1}\max_{|x|\le n}\mathbb{E}|G(\omega,x)|\le cn^{d+1}\lambda_n^{-1}\|G\|.$$

Taking $\lambda_n=n^{d+3}$ and using Borel–Cantelli gives (5.14). $\square$

Following [29] we introduce an orthogonal decomposition of the space $\overline{L}^2$. Set

$$\overline{L}^2_p=\mathrm{cl}\{\nabla F, F\in L^2\}\qquad \text{in } \mathcal{H},$$

and let $\overline{L}^2_s$ be the orthogonal complement of $\overline{L}^2_p$ in $\mathcal{H}$. (Here p stands for "potential" and s for "solenoidal.")

LEMMA 5.10. *Let $G\in\overline{L}^2_p$. Then for each x, $\mathbb{E}G(x,\omega)=0$.*

PROOF. Fix $x\in\mathbb{Z}^d$. Note first that if $G=\nabla F$, where $F\in L^2$, then $\mathbb{E}G(\omega,x)=\mathbb{E}(F_x-F)=\mathbb{E}F_x-\mathbb{E}F=0$.

Now let $G\in\overline{L}^2_p$. Then there exist $F_n\in L^2$ such that $G=\lim_n\nabla F_n$ in $\overline{L}^2$. Since $\mathbb{P}(Q_{0x}>0)=1$, it follows that $\nabla F_n(\omega,x)$ converges to $G(\omega,x)$ in $\mathbb{P}$-probability. By Lemma 5.9, for each $p\in[1,2)$ the sequence $\nabla F_n(\omega,x)$ is bounded in $L^p(\Omega,\mathbb{P})$, and therefore $\nabla F_n(\omega,x)$ converges to $G(\omega,x)$ in $L^1(\Omega,\mathbb{P})$. So $\mathbb{E}G(\omega,x)=\lim_n\mathbb{E}\nabla F_n(\omega,x)=0$. $\square$

We define the semi-direct product measure $\mathbb{P}^*=\mathbb{P}\times P^0_\omega$.

LEMMA 5.11. *Let $G\in\overline{L}^2_s$. Then*

$$\sum_{x\in\mathbb{Z}^d}Q_{0x}(\omega)G(\omega,x)=0,\qquad \mathbb{P}\text{-}a.s. \tag{5.16}$$

Hence $M_n=G(\omega,X_n)$ is a P^0_ω-martingale for $\mathbb{P}$-a.a. ω. Further, writing

$$\|G(\omega,\cdot)\|^2=\sum_x Q_{0,x}(\omega,x)G(\omega,x)^2,$$

we have

$$\langle M\rangle_n=\sum_{k=0}^{n-1}\|G(T_{\widehat{X}_k}\omega,\cdot)\|^2=\sum_{k=0}^{n-1}\|G(\widehat{Z}_k,\cdot)\|^2. \tag{5.17}$$

Hence

$$\mathbb{E}^*(M_n)^2=n\|G\|^2. \tag{5.18}$$

PROOF. If $F \in L^2$ and $G \in \overline{L}^2$, then using Lemma 5.5,

$$\begin{aligned}\sum_{x\in\mathbb{Z}^d} \mathbb{E} Q_{0x} G(\omega, x) F_x &= \sum_{x\in\mathbb{Z}^d} \mathbb{E} Q_{0x}(T_{-x}\omega) G(T_{-x}\omega, x) F_x(T_{-x}\omega) \\ &= \sum_{x\in\mathbb{Z}^d} \mathbb{E} Q_{0,-x}(\omega)\bigl(-G(\omega,-x)\bigr) F(\omega) \\ &= -\sum_{x\in\mathbb{Z}^d} \mathbb{E} Q_{0x}(\omega) G(\omega, x) F(\omega).\end{aligned}$$

Thus

$$\sum_{x\in\mathbb{Z}^d} \mathbb{E} Q_{0x} G(\cdot, x)(F + F_x) = 0. \tag{5.19}$$

If $G \in \overline{L}^2_s$, then

$$0 = \overline{\mathbb{E}}(G\nabla F) = \sum_x \mathbb{E} Q_{0x} G(\cdot, x)(F_x - F),$$

and so $\mathbb{E} \sum Q_{0x} G F = 0$. Since this holds for any $F \in L^2$ we obtain (5.16).

To show that M is a martingale it is enough to prove that for any x,

$$E^0_\omega\bigl(G(\omega, X_{n+1}) - G(\omega, X_n) | X_n = x\bigr) = 0. \tag{5.20}$$

However, using (5.16),

$$\begin{aligned}&E^0_\omega\bigl(G(\omega, X_{n+1}) - G(\omega, X_n) | X_n = x\bigr) \\ &\quad = \sum_y Q_{xy}(\omega)\bigl(G(\omega, y) - G(\omega, x)\bigr) \\ &\quad = \sum_y Q_{0,y-x}(T_x\omega) G(T_x\omega, y - x) = 0.\end{aligned}$$

Recall that $\langle M\rangle$ is the unique predictable process so that $M_n^2 - \langle M\rangle_n$ is a martingale. We have

$$\begin{aligned}E^x_\omega(M^2_{n+1} - M^2_n | \widehat{X}_n = y) &= E^x_\omega\bigl((M_{n+1} - M_n)^2 | \widehat{X}_n = y\bigr) \\ &= \sum_z Q_{yz}(\omega)\bigl(G(\omega, z) - G(\omega, y)\bigr)^2 \\ &= \sum_z Q_{0,z-y}(T_y\omega)\bigl(G(z - y, \omega)\bigr)^2 \\ &= \|G(T_y\omega, \cdot)\|^2,\end{aligned}$$

and (5.17) follows.

Finally,

$$\mathbb{E}^* M_n^2 = \mathbb{E}(E^0_\omega M_n^2) = \mathbb{E}(E^0_\omega \langle M\rangle_n) = \sum_{k=0}^{n-1} \mathbb{E}\|G(T_{\widehat{X}_k}\omega, \cdot)\|^2 = n\|G\|^2. \qquad \square$$

Let $\Pi : \mathbb{R}^d \to \mathbb{R}^d$ be the identity, and write Π_j for the jth coordinate of Π. Then $\Pi_j(y-x) = \Pi_j(y) - \Pi_j(x)$, so Π_j has the cocycle property. Further by (4.14),

$$\overline{\mathbb{E}}|\Pi_j|^2 = \mathbb{E}\sum_x Q_{0x}|x_j|^2 < \infty,$$

so $\Pi_j \in \mathcal{H}$. So we can define $\chi_j \in \overline{L}^2_p$ and $\Phi_j \in \overline{L}^2_s$ by

$$\Pi_j = \chi_j + \Phi_j \in \overline{L}^2_p \oplus \overline{L}^2_s;$$

this gives our definition of the corrector $\chi = (\chi_1, \dots, \chi_d) : \Omega \times \mathbb{Z}^d \to \mathbb{R}^d$. We will sometimes write $\chi(x)$ for $\chi(\cdot, x)$. Note that conventions about the sign of the corrector differ—compare [34] and [12]. As the environment process is invariant under isometries of $\mathbb{Z}^d$, $\|\Phi_j\| = \|\Phi_1\|$ for each $j = 1, \dots, d$.

The following proposition summarizes the properties of χ and Φ.

PROPOSITION 5.12. (a) $\widehat{M}_n = \widehat{X}_n - \chi(\omega, \widehat{X}_n)$ *is a* P^0_ω*-martingale.*

(b) *For each* $x \in \mathbb{Z}^d$, $\chi(\cdot, x) \in L^1$.

(c) *For each* $j = 1, \dots, d$

$$\mathbb{E}\sum_x Q_{0x}(\omega)|\Phi_j(\omega, x)|^2 = \|\Phi_1\|^2 < \infty.$$

(d) χ *is sublinear on average*; *for each* $\varepsilon > 0$

$$\lim_{n\to\infty} n^{-d} \sum_{|x|\le n} 1_{(|\chi(\omega,x)|>\varepsilon n)} = 0, \qquad \mathbb{P}\text{-}a.s. \tag{5.21}$$

PROOF. (a) and (b) are immediate from Lemmas 5.11 and 5.9, and (c) is immediate from the definition of Φ_j as a projection in $\overline{L}^2$. Let e_1 be the unit vector $e_1 = (1, 0, \dots, 0)$. By Lemma 5.10 we have $\mathbb{E}\chi(\cdot, e_1) = 0$. So since

$$\chi(\omega, ne_1) = \sum_{k=1}^n \chi\bigl(T_{(k-1)e_1}\omega, e_1\bigr) \tag{5.22}$$

and as χ has the cocycle property, the ergodic theorem implies that $\lim_n n^{-1}\chi(\omega, ne_1) = 0$ $\mathbb{P}$-a.s., and (d) then follows by the results in Section 6 of [26]. $\square$

LEMMA 5.13. *The processes* Z *and* $\widehat{Z}$ *are ergodic under the time shift on the environment space* Ω.

PROOF. This is well known; see [17], Lemma 4.9, and Section 3 of [10] for a careful proof in discrete time. $\square$

PROOF OF THEOREM 1.1. We begin with the VSRW. The arguments are very similar to those in [10, 12, 34], so we only mention the key points. We define

$$\widehat{M}_n = \Phi(\omega, \widehat{X}_n), \qquad \widehat{M}^{(\varepsilon)}_t = \varepsilon \widehat{M}_{\lfloor t/\varepsilon^2 \rfloor}, \qquad t \ge 0, \tag{5.23}$$

so that

$$\widehat{X}_t^{(\varepsilon)} = \varepsilon \widehat{X}_{\lfloor t/\varepsilon^2 \rfloor} = \widehat{M}_t^{(\varepsilon)} + \varepsilon \chi\big(\omega, \varepsilon^{-1} \widehat{X}_t^{(\varepsilon)}\big). \tag{5.24}$$

Thus it is sufficient to prove that the martingale $\widehat{M}^{(\varepsilon)}$ converges to a multiple of Brownian motion, and that for $\mathbb{P}$-a.a. ω, the second term in (5.24) converges in P_ω^0-probability to zero.

We start with the control of the corrector, and use [12], Theorem 2.4. This proves that if the corrector χ has polynomial growth, and is sublinear on average, then Gaussian upper bounds on the heat kernel imply pointwise sublinearity of χ. Thus, using (1.10), (5.14) and (5.21) we have that for $\mathbb{P}$-a.a. ω,

$$\lim_{n\to\infty} \max_{|x|\le n} \frac{|\chi(\omega, x)|}{n} = 0. \tag{5.25}$$

Given (5.25) the Gaussian upper bounds then imply that, for $\mathbb{P}$-a.a. ω,

$$\varepsilon \chi\big(\omega, \widehat{X}_{\lfloor t/\varepsilon^2 \rfloor}\big) \to 0 \qquad \text{in } P_\omega^0\text{-probability}. \tag{5.26}$$

For the convergence of $\widehat{M}^{(\varepsilon)}$, we proceed as in [10]. Let $v \in \mathbb{R}^d$ be a unit vector, write $\widehat{M}_n^v = v \cdot M_n$, and let

$$F_K(\omega) = E_\omega^0(|\widehat{M}_1^v|^2; |\widehat{M}_1^v| \ge K).$$

Then F_K is decreasing in K and

$$\mathbb{E} F_K \le \mathbb{E} F_0 \le d \|\Phi_1\|^2.$$

In the notation of Lemma 5.11, $F_0(\omega) = \|v \cdot \Phi(\omega, \cdot)\|^2$, and so by (5.17) the covariance process of $\widehat{M}^v$ is

$$\langle \widehat{M}^v \rangle_n = \sum_{k=0}^{n-1} F_0(\widehat{Z}_k).$$

So by Lemma 5.13 we have $n^{-1} \langle \widehat{M}^v \rangle_n \to \mathbb{E} F_0$, P_ω^0-a.s., for $\mathbb{P}$-a.a. ω.

Using the same arguments as in [10], Theorem 6.2, it is straightforward to check the conditions of the Lindeberg–Feller FCLT for martingales (see, e.g., [20], Theorem 7.7.3), and deduce that $v \cdot \widehat{M}^{(\varepsilon)}$ converges to a constant multiple of Brownian motion. Hence $\widehat{M}^{(\varepsilon)}$ converges to an $\mathbb{R}^d$-valued Brownian motion with nonrandom covariance matrix D given by $D_{ij} = \overline{\mathbb{E}} \Phi_i \Phi_j$. Since the law of the random variables $\omega(e)$ is invariant under symmetries of $\mathbb{Z}^d$, we deduce that there exists $\sigma_V^2 \ge 0$ such that $D = \sigma_V^2 I$, and that

$$\sigma_V^2 = \overline{\mathbb{E}} \Phi_1^2. \tag{5.27}$$

This establishes the convergence of $\widehat{X}^{(\varepsilon)}$; using Lemma 4.12 gives the convergence of $X^{(\varepsilon)}$ to the same limit.

The global upper bounds on $q_t^\omega(0,x)$ in Lemma 2.8 imply that if $\lambda > 0$ and $\lambda t^{1/2} \geq 1$, then

$$P_\omega^0(|X_t| \leq \lambda t^{1/2}) \leq ct^{-d/2}|B(0,\lambda t^{1/2})| \leq c'\lambda^d.$$

Hence there exists $\lambda > 0$ such that for all large t,

$$P_\omega^0(|X_t| > \lambda t^{1/2}) \geq \tfrac{1}{2},$$

which implies that $\sigma_V^2 > 0$.

We now consider the CSRW. Recall from (1.1) the definition of $\mu_x(\omega)$, set $F(\omega) = \mu_0(\omega)$, and

$$A_t = \int_0^t \mu_{X_s}\,ds = \int_0^t F(Z_s)\,ds. \tag{5.28}$$

Then if $\tau_t = \inf\{s \geq 0 : A_s \geq t\}$ is the inverse of A, the time changed process

$$Y_t = X_{\tau_t} \tag{5.29}$$

is the CSRW.

By the ergodic theorem for the process Z,

$$\lim_{t\to\infty} t^{-1}A_t = \mathbb{E}F = 2d\mathbb{E}\mu_e, \qquad \mathbb{P}^*\text{-a.s.}$$

So if $\mathbb{E}\mu_e < \infty$ then $\tau_t/t \to a$ a.s. where $a = 1/2d\mathbb{E}\mu_e > 0$. Let $Y_t^{(\varepsilon)} = \varepsilon Y_{t/\varepsilon^2}$. Then

$$Y_t^{(\varepsilon)} = X_{at}^{(\varepsilon)} + \big(Y_t^{(\varepsilon)} - X_{at}^{(\varepsilon)}\big) \tag{5.30}$$

and using Theorem 4.11 we have for any fixed $t_0 \geq 0$ that

$$\sup_{0\leq t\leq t_0} \big|Y_t^{(\varepsilon)} - X_{at}^{(\varepsilon)}\big| \tag{5.31}$$

converges in P_ω^0-probability to 0, for $\mathbb{P}$-a.a. ω. Thus $Y^{(\varepsilon)}$ converges to $\sigma_C B_t'$ where B' is a Brownian motion and $\sigma_C^2 = a\sigma_V^2 > 0$.

In the case when $\mathbb{E}\mu_e = \infty$ we have that $\tau_t/t \to 0$, and hence $Y^{(\varepsilon)}$ converges to a degenerate limit. □

We conclude this section by stating a local limit theorem for $q_t(x,y)$ (for the VSRW). Write

$$k_t(x) = (2\pi t\sigma_V^2)^{-d/2}e^{-|x|^2/2\sigma_V^2 t}$$

for the Gaussian heat kernel with diffusion constant σ_V^2 where σ_V^2 is as in Theorem 1.1.

THEOREM 5.14. *Let X be the VSRW. Let $T > 0$. For $x \in \mathbb{R}^d$ write $\lfloor x \rfloor = (\lfloor x_1 \rfloor, \dots, \lfloor x_d \rfloor)$. Then*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}^d} \sup_{t\geq T} |n^{d/2} q^{\omega}_{nt}(0, \lfloor n^{1/2} x \rfloor) - k_t(x)| = 0, \qquad \mathbb{P}\text{-a.s.} \tag{5.32}$$

PROOF. This is proved as in Section 4 of [3]. We have to verify Assumptions 4.1 and 4.4 in [3], but this is straightforward given the invariance principle and heat kernel bounds in Theorems 1.1 and 1.2, and the PHI in Theorem 4.7.

Note that as ν is the invariant measure for X, in Assumption 4.1(d) all we need is that $\nu(\Lambda_n(x,r))/(2n^{1/2}r)d$ converges, and as $\nu_x = 1$ for all x; this is easy. (Here $\Lambda_n(x,r) = (xn^{1/2} + [-rn^{1/2}, rn^{1/2}]^d) \cap \mathbb{Z}^d$.) □

REMARK 5.15. In this section we have constructed a corrector $\chi(\omega, x)$ so that the process

$$M_n = X_n - \chi(\omega, X_n), \qquad n \in \mathbb{Z}_+, \tag{5.33}$$

is a (discrete time) martingale. It is natural to ask if $\chi(\omega, \cdot)$ is also a corrector for the continuous time process X_t.

For the RCM with i.i.d. conductances it is straightforward to check that the condition in Lemma 2.21 involving the set $A(K)$ holds $\mathbb{P}$-a.s. (see [12], Lemma 3.1, for a similar argument). We can therefore use Lemma 2.21 with $h(\cdot) = \chi(\omega, \cdot)$ to deduce that

$$M_t = X_t - \chi(\omega, X_t), \qquad t \in \mathbb{R}_+, \tag{5.34}$$

is, for $\mathbb{P}$-a.a. ω, a P^0_ω-martingale.

6. General ergodic environments. We conclude this paper with some remarks on more general ergodic random environments. First, note that the proof of the invariance principle in Section 5 just uses the facts that the environment is stationary, symmetric and ergodic, and that the heat kernel bounds in Theorem 1.2 hold.

In the proof of Theorem 1.2 the full strength of the assumption that μ_e were i.i.d. was only used at one point, in Theorem 4.3, where we controlled the probability that a ball was not very good. The heat kernel upper bounds in Section 2 only require Assumptions 2.1 and 2.6, together with a comparison of the metrics $\widetilde{d}(x, y)$ and $d(x, y)$. Given these upper bounds, and using the fact that $\mu_e \geq 1$, no additional hypotheses on μ_e were needed to obtain the lower bounds in Section 3. We therefore have the following:

THEOREM 6.1. *Let μ_e, $e \in E_d$ be a stationary symmetric ergodic environment, satisfying for some $c_1 > 0$,*

$$\mu_e \in [c_1, \infty) \qquad \textit{for all } e \in E_d, \mathbb{P}\text{-a.s.} \tag{6.1}$$

Let $\widetilde{d}_\omega(x, y)$ *be the metric given by the first passage percolation construction of* (2.6), *and* (*as in Definition* 2.9) *let* $V_x(\lambda)$ *be the smallest integer such that* $(x, V_x(\lambda))$ *is* λ*-very good. Suppose that there exists* $\lambda_0 < \infty$ *and* $\eta \in (0, 1)$ *such that*

$$\mathbb{P}\big(V_x(\lambda_0) \geq n\big) \leq c_1 e^{-c_2 n^\eta}. \tag{6.2}$$

Then the conclusions of Theorems 1.1, 1.2(a)–(c), 1.3, 4.7 *and* 5.14 *all hold for the environment* (μ_e).

PROOF. We begin by considering the heat kernel bounds in Theorem 1.2. As in Lemma 4.1, it is immediate that Assumptions 2.1 and 2.6 hold for (μ_e), $\mathbb{P}$-a.s. Using the hypothesis (6.2) instead of Theorem 4.3, the arguments in Section 4 [except for Theorem 4.9(a), for which see Remark 6.2 below] hold in this more general context, and give Theorems 1.2 and 4.7.

Given Theorem 1.2, the arguments in Section 5 then give the invariance principle (Theorem 1.1) and local limit theorem (Theorem 5.14).

Combining these results gives the Green function estimates in Theorem 1.3. □

REMARK 6.2. The proof of Theorem 4.9 used the fact that the bounds in Theorem 4.6 hold for $1 - \eta$ as well as for η. If we only have (6.2) then we obtain

$$\mathbb{E}q_t^\omega(x, y) \leq c_1 t^{-d/2} e^{-c_2|x-y|^2/t} \qquad \text{if } t \geq c_3 \vee |x - y|^{1+\eta}, \tag{6.3}$$

$$\mathbb{E}q_t^\omega(x, y) \geq c_4 t^{-d/2} e^{-c_5|x-y|^2/t} \qquad \text{if } t \geq c_6 \vee |x - y|^{2-\eta}. \tag{6.4}$$

REMARK 6.3. If μ_e is bounded and bounded away from 0, so there exist $0 < c_1 \leq c_2 < \infty$ such that $\mathbb{P}(\mu_e \in [c_1, c_2]) = 1$, then the metrics $d(x, y)$ and $\widetilde{d}(x, y)$ are comparable. So, taking λ_0 large enough, (6.2) holds.

REMARK 6.4. See [12], Lemma 3.1, or [30], Lemma 5.3, for percolation arguments which are more robust than Theorem 4.3 and which may be useful for establishing (6.2) in more general contexts.

REMARK 6.5. If μ_e is stationary and ergodic, but not invariant with respect to symmetries of $\mathbb{Z}^d$, then if (6.2) holds, we still obtain Theorem 1.2, and the convergence of $X^{(n)}$ to a Brownian motion with covariance matrix D. However, D need not be diagonal.

REMARK 6.6. Unlike ergodic bounded conductance models, the results of this paper certainly do not hold for all unbounded stationary symmetric ergodic random environments. For example, let $d = 2, 3, 4$ and let $\mathcal{T}$ be a uniform spanning tree on $\mathbb{Z}^d$ (see [9]). Then $\mathcal{T}$ is 1-sided, so from each $x \in \mathbb{Z}^d$ there is a unique self-avoiding path γ_x to infinity. Let $a(x)$ be the first point on this path. Then $a : \mathbb{Z}^d \to \mathbb{Z}^d$ and the path γ_x is $\{x, a(x), a^2(x), \ldots\}$.

Let $N(x)$ be the set of points in $\mathcal{T}$ which are disconnected from infinity by deleting the bond $\{x, a(x)\}$, and let $n(x) = |N(x)|$. As $x \in N(x)$, $n(x) \geq 1$ for all x. Let $\mu_e = 1$ for edges $e \in E_d$ which are not in $\mathcal{T}$. Each edge $e \in \mathcal{T}$ is of the form $e = \{x, a(x)\}$ for some x, set

$$\mu_{\{x,a(x)\}} = n(x)e^{n(x)^2}.$$

Let T_i, $i \geq 1$, be the jump times of the VSRW X. Then

$$P_\omega^x\big(X_{T_1} \neq a(x)\big) = \frac{\sum_{y \neq a(x)} \mu_{xy}}{\mu_{x,a(x)} + \sum_{y \neq a(x)} \mu_{xy}}. \tag{6.5}$$

Fix x, and let the neighbors of x in $\mathcal{T}$ be $a(x), y_1, \dots, y_k$. Then

$$\sum_{y \neq a(x)} \mu_{xy} = (2d - k - 1) + \sum_{i=1}^{k} n(y_i)e^{n(y_i)^2}.$$

Since $\mu_{x,a(x)} = n(x)e^{n(x)^2}$ and $n(x) = 1 + \sum n(y_i)$, it is easy to see that

$$p_0(x) = P_\omega^x\big(X_{T_1} \neq a(x)\big) \leq 2de^{-n(x)^2} + \max_i e^{n(y_i)^2 - n(x)^2} \leq 2de^{-n(x)^2} + e^{-n(x)/d}.$$

So $\sum_k p_0(a^k(x)) < \infty$, and it follows that ultimately the process X moves to infinity along a path γ_x for some x. Since $\sum_k \mu^{-1}_{a^k(x),a^{k+1}(x)} < \infty$ this takes finite time. Hence the quenched invariance principle Theorem 1.1 fails, as well as the Gaussian bounds in Theorem 1.2.

Acknowledgments. The authors thank J. Černý and T. Kumagai for valuable discussions.

REFERENCES

[1] BARLOW, M. T. (2004). Random walks on supercritical percolation clusters. *Ann. Probab.* **32** 3024–3084. MR2094438

[2] BARLOW, M. T. and ČERNÝ, J. (2009). Convergence to fractional kinetics for random walks associated with unbounded conductances. Preprint.

[3] BARLOW, M. T. and HAMBLY, B. M. (2009). Parabolic Harnack inequality and local limit theorem for percolation clusters. *Electron. J. Probab.* **14** 1–26. MR2471657

[4] BASS, R. F. (2002). On Aronson's upper bounds for heat kernels. *Bull. Lond. Math. Soc.* **34** 415–419. MR1897420

[5] BARLOW, M. T. and BASS, R. F. (1989). The construction of Brownian motion on the Sierpiński carpet. *Ann. Inst. H. Poincaré Probab. Statist.* **25** 225–257. MR1023950

[6] BEN AROUS, G., ČERNÝ, J. and MOUNTFORD, T. (2006). Aging in two-dimensional Bouchaud's model. *Probab. Theory Related Fields* **134** 1–43. MR2221784

[7] BEN AROUS, G. and ČERNÝ, J. (2007). Scaling limit for trap models on $\mathbb{Z}^d$. *Ann. Probab.* **35** 2356–2384. MR2353391

[8] BEN AROUS, G. and ČERNÝ, J. (2008). The arcsine law as a universal aging scheme for trap models. *Comm. Pure Appl. Math.* **61** 289–329. MR2376843

[9] BENJAMINI, I., LYONS, R., PERES, Y. and SCHRAMM, O. (2001). Uniform spanning forests. *Ann. Probab.* **29** 1–65. MR1825141

[10] BERGER, N. and BISKUP, M. (2007). Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Related Fields* **137** 83–120. MR2278453

[11] BERGER, N., BISKUP, M., HOFFMAN, C. E. and KOZMA, G. (2008). Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann. Inst. H. Poincaré Probab. Statist.* **44** 374–392. MR2446329

[12] BISKUP, M. and PRESCOTT, T. M. (2007). Functional CLT for random walk among bounded random conductances. *Electron. J. Probab.* **12** 1323–1348. MR2354160

[13] BOIVIN, D. and DERRIENNIC, Y. (1991). The ergodic theorem for additive cocycles of $\mathbf{Z}^d$ or $\mathbf{R}^d$. *Ergodic Theory Dynam. Systems* **11** 19–39. MR1101082

[14] CARLEN, E. A., KUSUOKA, S. and STROOCK, D. W. (1987). Upper bounds for symmetric Markov transition functions. *Ann. Inst. H. Poincaré Probab. Statist.* **23** 245–287. MR898496

[15] DAVIES, E. B. (1993). Large deviations for heat kernels on graphs. *J. Lond. Math. Soc.* (2) **47** 65–72. MR1200978

[16] DAVIES, E. B. (1993). Analysis on graphs and noncommutative geometry. *J. Funct. Anal.* **111** 398–430. MR1203460

[17] DE MASI, A., FERRARI, P. A., GOLDSTEIN, S. and WICK, W. D. (1989). An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Stat. Phys.* **55** 787–855. MR1003538

[18] DELMOTTE, T. (1999). Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoamericana* **15** 181–232. MR1681641

[19] DELMOTTE, T. and DEUSCHEL, J.-D. (2005). On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to $\nabla\phi$ interface model. *Probab. Theory Related Fields* **133** 358–390. MR2198017

[20] DURRETT, R. (1996). *Probability: Theory and Examples*, 3rd ed. Duxbury Press, Belmont, CA. MR1609153

[21] FABES, E. B. and STROOCK, D. W. (1986). A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Ration. Mech. Anal.* **96** 327–338. MR855753

[22] FUKUSHIMA, M., OSHIMA, Y. and TAKEDA, M. (1994). *Dirichlet Forms and Symmetric Markov Processes. De Gruyter Studies in Mathematics* **19**. De Gruyter, Berlin. MR1303354

[23] GRIGOR'YAN, A. A. (1991). The heat equation on noncompact Riemannian manifolds. *Mat. Sb.* **72** 47–77. MR1098839

[24] GRIGOR'YAN, A. (1997). Gaussian upper bounds for the heat kernel on arbitrary manifolds. *J. Differential Geom.* **45** 33–52. MR1443330

[25] KESTEN, H. (1986). Aspects of first passage percolation. In *École D'été de Probabilités de Saint–Flour, XIV—1984. Lecture Notes in Math.* **1180** 125–264. Springer, Berlin. MR876084

[26] KEYNES, H. B., MARKLEY, N. G. and SEARS, M. (1995). Ergodic averages and integrals of cocycles. *Acta Math. Univ. Comenian. (N.S.)* **64** 123–139. MR1360992

[27] KIPNIS, C. and VARADHAN, S. R. S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** 1–19. MR834478

[28] KOZLOV, S. (1985). The method of averaging and walks in inhomogeneous environments. *Russian Math. Surveys* **40** 73–145.

[29] MATHIEU, P. and PIATNITSKI, A. (2007). Quenched invariance principles for random walks on percolation clusters. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **463** 2287–2307. MR2345229

[30] MATHIEU, P. (2008). Quenched invariance principles for random walks with random conductances. *J. Stat. Phys.* **130** 1025–1046. MR2384074

[31] MOURRAT, J.-C. (2009). Variance decay for functionals of the environment viewed by the particle. Preprint.

[32] NASH, J. (1958). Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.* **80** 931–954. MR0100158

[33] SALOFF-COSTE, L. (1992). A note on Poincaré, Sobolev, and Harnack inequalities. *Int. Math. Res. Not. IMRN* **2** 27–38. MR1150597

[34] SIDORAVICIUS, V. and SZNITMAN, A.-S. (2004). Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Related Fields* **129** 219–244. MR2063376

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, BC V6T 1Z2
CANADA
E-MAIL: barlow@math.ubc.ca

FACHBEREICH MATHEMATIK
TECHNISCHE UNIVERSITÄT BERLIN
STRASSE DES 17. JUNI 136
D-10623 BERLIN
GERMANY
E-MAIL: deuschel@math.tu-berlin.de