Commun. Math. Phys. 263, 21–64 (2006)
Digital Object Identifier (DOI) 10.1007/s00220-005-1492-5

Communications in
Mathematical Physics

Mott Law as Lower Bound for a Random Walk in a Random Environment

A. Faggionato[1], **H. Schulz-Baldes**[2], **D. Spehner**[3]

[1] Weierstrass Institut für Angewandte Analysis und Stochastic, 10117 Berlin, Germany
[2] Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany
[3] Fachbereich Physik, Universität Duisburg-Essen, 45117 Essen, Germany

Received: 21 July 2004 / Accepted: 16 September 2005
Published online: 24 January 2006 – © Springer-Verlag 2006

Abstract: We consider a random walk on the support of an ergodic stationary simple point process on $\mathbb{R}^d$, $d \geq 2$, which satisfies a mixing condition w.r.t. the translations or has a strictly positive density uniformly on large enough cubes. Furthermore the point process is furnished with independent random bounded energy marks. The transition rates of the random walk decay exponentially in the jump distances and depend on the energies through a factor of the Boltzmann-type. This is an effective model for the phonon-induced hopping of electrons in disordered solids within the regime of strong Anderson localization. We show that the rescaled random walk converges to a Brownian motion whose diffusion coefficient is bounded below by Mott's law for the variable range hopping conductivity at zero frequency. The proof of the lower bound involves estimates for the supercritical regime of an associated site percolation problem.

1. Introduction

1.1. Main Result. Let us directly describe the model and the main results of this work, deferring a discussion of the underlying physics to the next section. Suppose given an infinite countable set of random points $\{x_j\} \subset \mathbb{R}^d$ distributed according to some ergodic stationary simple point process. One can identify this set with the simple counting measure $\hat{\xi} = \sum_j \delta_{x_j}$ having $\{x_j\}$ as its support, and then write $x \in \hat{\xi}$ if $x \in \{x_j\}$. The σ-algebra $\mathcal{B}(\hat{\mathcal{N}})$ on the space $\hat{\mathcal{N}}$ of counting measures on $\mathbb{R}^d$ is generated by the family of subsets $\{\hat{\xi} \in \hat{\mathcal{N}} : \hat{\xi}(B) = n\}$, where $B \subset \mathbb{R}^d$ is Borel and $n \in \mathbb{N}$. The distribution $\hat{\mathcal{P}}$ of the point process is a probability on the measure space $(\hat{\mathcal{N}}, \mathcal{B}(\hat{\mathcal{N}}))$. It is stationary and ergodic w.r.t. the translations $x \mapsto x + y$ of $\mathbb{R}^d$. In the sequel, we need to impose boundedness of some κth moment defined by

$$\rho_\kappa := \mathbf{E}_{\hat{\mathcal{P}}}\big(\hat{\xi}(C_1)^\kappa\big), \tag{1}$$

where $C_1 = [-\frac{1}{2}, \frac{1}{2}]^d$ and $\mathbf{E}_{\hat{\mathcal{P}}}$ is the expectation w.r.t. $\hat{\mathcal{P}}$. Then $\rho = \rho_1$ is the so–called intensity of the process.

To each x_j is associated a random energy mark $E_{x_j} \in [-1, 1]$. These marks are drawn independently and identically according to a probability measure ν. Again, $\{(x_j, E_{x_j})\}$ is naturally identified with an element ξ of the space $\mathcal{N}$ of counting measures on $\mathbb{R}^d \times [-1, 1]$, and the distribution $\mathcal{P}$ of the marked process is a measure on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$ (with $\mathcal{B}(\mathcal{N})$ defined similarly to $\mathcal{B}(\hat{\mathcal{N}})$). The distribution $\mathcal{P}$ is said to be the *ν–randomization* of $\hat{\mathcal{P}}$ [Kal]. It is stationary and ergodic w.r.t. $\mathbb{R}^d$–translations. In order to assure that $\{x_j\}$ contains the origin, we consider the measurable subset $\mathcal{N}_0 = \{\xi \in \mathcal{N} : \xi(\{0\} \times [-1, 1]) = 1\}$ furnished with the σ-algebra $\mathcal{B}(\mathcal{N}_0) = \{A \cap \mathcal{N}_0 : A \in \mathcal{B}(\mathcal{N})\}$. The *random environment* is given by a configuration $\xi \in \mathcal{N}_0$ randomly chosen along the Palm distribution $\mathcal{P}_0$ associated to $\mathcal{P}$. Roughly, one can think of $\mathcal{P}_0$ as the probability on $(\mathcal{N}_0, \mathcal{B}(\mathcal{N}_0))$ obtained by conditioning $\mathcal{P}$ to the event $\mathcal{N}_0$ (see Sect. 2). Note that almost each environment is a simple counting measure, and therefore it can be identified with its support as we will do in what follows.

For a fixed environment $\xi \equiv \{(x_j, E_{x_j})\} \in \mathcal{N}_0$ let us consider a continuous-time random walk over the points $\{x_j\}$ starting at the origin $x = 0$ with transition rates from $x \in \hat{\xi}$ to $y \in \hat{\xi}$ given by

$$c_{x,y}(\xi) := \exp\big(-|x-y| - \beta(|E_x - E_y| + |E_x| + |E_y|)\big), \ x \neq y\,, \tag{2}$$

where $\beta > 0$ is the inverse temperature. More precisely, let $\Omega_\xi = D([0, \infty), \mathrm{supp}(\hat{\xi}))$ be the space of right-continuous paths on the support of $\hat{\xi}$ having left limits, endowed with the Skorohod topology [Bil]. Let us write $(X_t^\xi)_{t\geq 0}$ for a generic element of Ω_ξ. If $\mathbf{P}_0^\xi$ denotes the distribution on $(\Omega_\xi, \mathcal{B}(\Omega_\xi))$ of the above random walk starting at the origin, then the set of stationary transition probabilities $p_t^\xi(y|x) := \mathbf{P}_0^\xi(X_{s+t}^\xi = y | X_s^\xi = x)$, $x, y \in \hat{\xi}$, $t \geq 0$, $s > 0$ satisfy the following conditions for small values of t [Bre]:

(C1) $p_t^\xi(y|x) = c_{x,y}(\xi)\, t + o(t)$ if $x \neq y$;

(C2) $p_t^\xi(x|x) = 1 - \lambda_x(\xi)\, t + o(t)$ with $\lambda_x(\xi) := \sum_{y\in\hat{\xi}} c_{x,y}(\xi)$, where $c_{x,x}(\xi) := 0$.

It is verified in Appendix A that, provided that $\rho_2 < \infty$, no explosions occur and thus the random walk is well-defined for $\mathcal{P}_0$–almost all ξ.

Our main interest concerns the long time asymptotics of the random walk and the diffusion matrix D defined by

$$(a \cdot Da) = \lim_{t\to\infty} \frac{1}{t}\, \mathbf{E}_{\mathcal{P}_0}\Big(\mathbf{E}_{\mathbf{P}_0^\xi}\big((X_t^\xi \cdot a)^2\big)\Big), \qquad a \in \mathbb{R}^d\,, \tag{3}$$

where $(a \cdot b)$ denotes the scalar product of the vectors a and b in $\mathbb{R}^d$. The main results of the work are (i) the existence of the limit (3) in any dimension $d \geq 1$ as well as the convergence of the (diffusively rescaled) random walk to a Brownian motion with finite covariance matrix $D \geq 0$; (ii) a quantitative lower bound on D in dimension $d \geq 2$ under given assumptions on the energy distribution ν and either one of the following two technical hypotheses. Let ℓ denote the Lebesgue measure and $C_N = [-N/2, N/2]^d$. Given $A \subset \mathbb{R}^d$, let $\mathcal{F}_A$ be the σ–subalgebra in $\mathcal{B}(\hat{\mathcal{N}})$ generated by the random variables $\hat{\xi}(B)$ with $B \subset A$ and $B \in \mathcal{B}(\mathbb{R}^d)$.

(H1) $\hat{\mathcal{P}}$ admits a lower bound $\rho' > 0$ on the point density:

$$\hat{\xi}(C_N) \geq \rho'\, \ell(C_N), \qquad \forall\ N \geq N_0\,,\ \hat{\mathcal{P}}\text{-a.s.}\,, \tag{4}$$

with ρ' and N_0 independent on $\hat{\xi}$.

(H2) $\hat{\mathcal{P}}$ satisfies the following mixing condition: there exists a function $h : \mathbb{R}_+ \to \mathbb{R}_+$ with $h(r) \leq c(1 + r^{2d+7+\delta})^{-1}$ for some $c, \delta > 0$ such that for any $r_2 \geq r_1 > 1$,

$$\left|\hat{\mathcal{P}}(A|\mathcal{F}_{\mathbb{R}^d \setminus C_{r_2}}) - \hat{\mathcal{P}}(A)\right| \leq r_1^d \, r_2^{d-1} \, h(r_2 - r_1) \,, \qquad \forall \, A \in \mathcal{F}_{C_{r_1}} \,, \hat{\mathcal{P}}\text{-a.s.} \quad (5)$$

We feel that Hypotheses (H1) and (H2) cover nearly all interesting examples see however Example 2 below. The uniform lower bound (H1) holds in the case of random and quasiperiodic tilings and, more generally, the so-called Delone sets [BHZ]. The type of mixing condition (H2) is inspired by decorrelation estimates holding for Gibbs measures of spin systems in a high temperature phase [Mar]. It is satisfied for a stationary Poisson point process as well as for point processes with finite range correlations. Due to the stationarity of $\hat{\mathcal{P}}$, (H2) implies that $\hat{\mathcal{P}}$ is a mixing, and, in particular, ergodic point process (see [DV, Chap. 10]). We can now state more precisely the above-mentioned results.

Theorem 1. *Let $\hat{\mathcal{P}}$ be the distribution of an ergodic stationary simple point process on $\mathbb{R}^d$, let $\mathcal{P}$ be the distribution of its ν–randomization with a probability measure ν on $[-1, 1]$, and let $\mathcal{P}_0$ be the Palm distribution associated to $\mathcal{P}$. Assume that $\rho_{12} < \infty$ and that*

$$\xi = \sum_j \delta_{(x_j, E_j)} \quad \Rightarrow \quad \xi \neq S_x \xi := \sum_j \delta_{(x_j - x, E_j)} \quad \forall \, x \in \mathbb{R}^d \setminus \{0\}, \quad \mathcal{P} \text{ a.s.} \quad (6)$$

Condition (6) *is automatically satisfied if ν is not a Dirac measure. Then:*

(i) The limit in (3) *exists and the rescaled process $\underline{Y}^{\xi,\varepsilon} = (\varepsilon X^{\xi}_{t\varepsilon^{-2}})_{t \geq 0}$ defined on $(\Omega_\xi, \mathbf{P}_0^\xi)$ converges weakly in $\mathcal{P}_0$-probability as $\varepsilon \to 0$ to a Brownian motion $\underline{W}_D$ with covariance matrix D. Namely, for any bounded continuous function F on the path space $D([0, \infty), \mathbb{R}^d)$ endowed with the Skorohod topology,*

$$\mathbf{E}_{\mathbf{P}_0^\xi}\Big(F\big(\underline{Y}^{\xi,\varepsilon}\big)\Big) \to \mathbf{E}\Big(F\big(\underline{W}_D\big)\Big) \qquad \text{in } \mathcal{P}_0\text{-probability} \,.$$

(ii) Suppose $d \geq 2$ and let either (H1) *or* (H2) *be satisfied. Furthermore, suppose that there are some positive constants α, c_0 such that, for any $0 < E \leq 1$,*

$$\nu([-E, E]) \geq c_0 \, E^{1+\alpha} \,. \quad (7)$$

Then

$$D \geq c_1 \, \beta^{-\frac{d(\alpha+1)}{\alpha+1+d}} \, \exp\left(-c_2 \, \beta^{\frac{\alpha+1}{\alpha+1+d}}\right) \mathbf{1}_d \,, \quad (8)$$

where $\mathbf{1}_d$ is the $d \times d$ identity matrix and c_1 and c_2 are some positive β-independent constants.

The important factor in the lower bound (8) is the exponential factor and not the power law in front of it (on which we comment below though). Based on the following heuristics due to Mott [Mot, SE], we expect that the expression in the exponential in (8) captures the good asymptotic behavior of $\ln D$ in the low temperature limit $\beta \uparrow \infty$ if $\nu([-E, E]) \sim c_0 E^{1+\alpha}$ as $E \downarrow 0$. Indeed, as β becomes larger, the rates (2) fluctuate widely with (x, y) because of the exponential energy factor. The low temperature limit effectively selects only jumps between points with energies in a small interval $[-E(\beta), E(\beta)]$ shrinking to zero as $\beta \uparrow \infty$. Assuming that D is determined by those

jumps with the largest rate, one obtains directly the characteristic exponential factor on the r.h.s. of (8) by maximizing these rates for a fixed temperature under the constraint that the mean density of points x_j with energies in $[-E(\beta), E(\beta)]$ is equal to $\rho\,\nu([-E(\beta), E(\beta)]) \sim c_0 \rho E(\beta)^{1+\alpha}$. One speaks of *variable range hopping* since the characteristic mean distance $|x - y|$ between sites with optimal jump rates varies heavily with the temperature. A crucial (and physically reasonable, as discussed below) element of this argument is the independence of the energies E_x. The selection of the points $\{x_j\}$ with energies in the window $[-E(\beta), E(\beta)]$ then corresponds mathematically to a p-thinning with $p = \nu([-E(\beta), E(\beta)])$. It is a well-known fact (see *e.g.* [Kal, Theorem 16.19]) that an adequate rescaling of the p-thinning of a stationary point process converges in the limit $p \downarrow 0$ (corresponding here to $\beta \uparrow \infty$) to a stationary Poisson point process (PPP). Hence one might call the stationary PPP the normal form of a model leading Mott's law, namely the exponential factor on the r.h.s. of (8), and we believe that proving the upper bound corresponding to (8) should therefore be most simple for the PPP. In dimension $d = 1$, a different behavior of D is expected [LB] and this will not be considered here. Note that statement (i) does not necessarily imply that the motion of the particle is diffusive at large time, since it could happen that $D = 0$.

The preexponential factor in (8) can be improved to $\beta^{\frac{(\alpha+1)(2-d)}{(1+\alpha+d)}}$ by means of formal scaling arguments on the formulas in Sects. 4 to 6. As we are not sure that this is optimal and we do not control the constant c_2 in (8) anyway, we choose not to develop this improvement in detail.

1.2. Physical discussion. Our main motivation for studying the above model comes from its importance for *phonon-assisted hopping conduction* [SE] in disordered solids in which the Fermi level (set equal to 0 above) lies in a region of *strong Anderson localization*. This means that the electron Hamiltonian has exponentially localized quantum eigenstates with localization centers x_j if the corresponding energies E_{x_j} are close to the Fermi level. The DC conductivity of such materials would vanish if it were not for the lattice vibrations (phonons) at nonzero temperature. They induce transitions between the localized eigenstates, the rate of which can be calculated from first principle by means of the Fermi golden rule [MA, SE]. In the variable range hopping regime at low temperature, the Markov and adiabatic (or rotating wave) approximations can be used to treat quantum mechanically the electron-phonon coupling [Spe]. Coherences between electronic eigenstates with different energies decay very rapidly under the resulting dissipative electronic dynamics and one can show that the hopping DC conductivity of the disordered solid coincides with the conductivity associated with a Markov jump process on the set of localization centers $\{x_j\}$, hence justifying the use of a model of classical mechanics [BRSW]. Because Pauli blocking due to Fermi statistics of the electrons has to be taken into account, this leads to a rather complicated exclusion process (*e.g.* [Qua, FM]). If, however, the blocking is treated in an *effective medium* (or mean field) approximation, one obtains a family of independent random walks with rates which are given by (2) in the limit $\beta \uparrow \infty$ [MA, AHL].

Let us discuss the remaining aspects of the model. The stationarity of the underlying simple point process $\{x_j\}$ simply reflects that the material is homogeneous, while the independence of the energy marks is compatible with Poisson level statistics, which is a general rough indicator for the localization regime and has been proven to hold for an Anderson model [Min]. The exponent α allows to model a possible Coulomb pseudogap in the density of states [SE].

Having in mind the Einstein relation between the conductivity and the diffusion coefficient (which can be stated as a theorem for a number of models [Spo]), the lower bound (8) gives a lower bound on the hopping DC condutivity. In the above materials, the DC conductivity shows experimentally Mott's law, namely a low-temperature behavior which is well approximated by the exponential factor in the r.h.s. of (8) with $\alpha = 0$, as predicted by Mott [Mot] based on the optimization argument discussed above. In certain materials having a Coulomb pseudogap in the density of state, Mott's law with $\alpha = d-1$ is observed, as predicted by Efros and Shklovskii [EF]. A first convincing justification of Mott's argument was given by Ambegoakar, Halperin and Langer [AHL], who first reduced the hopping model to a related random resistor network, in a manner similar to the work of Miller and Abrahams [MA], and then pointed out that the constant c_2 in (8) can be estimated using percolation theory [SE]. Our proof of the lower bound (8) is inspired by this work. Let us also mention that the low frequency AC conductivity (response to an oscillating electric field) in disordered solids has recently been studied within a quantum-mechanical one-body approximation in [KLP]. Here the energy necessary for a jump between localized states comes from a resonance at the frequency of the external electric field rather than a phonon. It leads to another well-known formula for the conductivity which is also due to Mott.

1.3. Overview. Let us develop the main ideas of the proof of Theorem 1, leaving precise statements and their proofs to the following sections. The model described above is a *random walk in a random environment*. A main tool used in this work is the contribution of De Masi, Ferrari, Goldstein and Wick [DFGW] which is based on prior work by Kipnis and Varadhan [KV]. They construct a new Markov process, called the *environment viewed from the particle*, which allows to translate the homogeneity of the medium into properties of the random walk. In Sect. 3, we argue that X_t^ξ has finite moments w.r.t. $\int \mathcal{P}_0(d\xi)\mathbf{P}_0^\xi$ (Proposition 1) and study the generator of the process environment viewed from the particle when the initial environment is chosen according to the Palm distribution $\mathcal{P}_0$ (Propositions 2 and 3), thus allowing to apply the general Theorem 2.2 of [DFGW] to deduce the existence of the limit (3). The convergence to a Brownian motion stated in Theorem 1 also follows, but this could have been obtained (avoiding an analysis of the infinitesimal generator) by applying Theorem 17.1 of [Bil] and Theorem 2.1 of [DFGW]. The results of [DFGW] also lead to a variational formula for the diffusion matrix D (Theorem 2 below). The main virtue of this formula is that it allows to bound D from below through bounds on the transition rates.

The first step in proving Theorem 1(ii) is to define a new random walk with transition rates bounded above by the rates (2). This is done in Sect. 4 in the following way. For a fixed configuration $\xi \in \mathcal{N}_0$ of the environment, consider the set $\{x_j^c\} = \{x_j : |E_{x_j}| \leq E_c\}$ of all random points having energies inside a given energy window $[-E_c, E_c]$ with $0 < E_c \leq 1$. The distribution $\hat{\mathcal{P}}^c$ of these points is obtained from $\mathcal{P}$ by a δ_c-thinning with $\delta_c = \nu([-E_c, E_c])$. Given a cut-off distance $r_c > 0$, consider the random walk on $\mathrm{supp}(\hat{\xi})$ with the transition rates $\hat{c}_{x,y}(\xi) = \chi(|x-y| \leq r_c)\chi(|E_x| \leq E_c)\chi(|E_y| \leq E_c)$, where χ is the characteristic function. Since we want this random walk to have a strictly positive diffusion coefficient in the limit $E_c \to 0$, one must choose r_c such that the mean number of points x_j with energies in $[-E_c, E_c]$ inside a ball of radius r_c is larger than an E_c-independent constant $c_3 > 0$. This mean number is equal to $c_4 \delta_c r_c^d$ and is larger than $c_5 E_c^{1+\alpha} r_c^d$ by assumption (7), where c_4 and c_5 are constants depending on ρ and d only. Hence $r_c = c_6\, E_c^{-(1+\alpha)/d}$. It is shown in Proposition 5 that the diffusion matrix of

the new random walk is equal to $\delta_c\, D(r_c, E_c)$, where $D(r_c, E_c)$ is the diffusion matrix of a random walk on $\{x_j^c\}$ with energy-independent transition rates $\chi(|x-y| \le r_c)$. By the monotonicity of D in the jump rates and since $c_{x,y}(\xi) \ge \exp(-r_c - 4\beta E_c)\, \hat{c}_{x,y}(\xi)$, one gets using also assumption (7) and the constant c_0 therein:

$$D \;\ge\; c_0\, E_c^{1+\alpha}\, e^{-r_c - 4\beta E_c}\, D(r_c, E_c)\;. \tag{9}$$

In Sect. 5, a lower bound on $D(r_c, E_c)$ is obtained by considering periodic approximants (in the limit of large periods) as in [DFGW]. The diffusion coefficient of these approximants can be computed as the resistance of a random resistor network. The resistance of the random resistor network is bounded by invoking estimates from percolation theory in Sect. 6, hence showing that, if r_c is large enough, $D(r_c, E_c) > c_7\, \mathbf{1}_d$, where $c_7 > 0$ is independent on E_c, β. Recalling that $r_c = c_6\, E_c^{-(1+\alpha)/d}$, an optimization w.r.t. E_c of the right member of (9) then yields $E_c = c_8 \beta^{-\frac{d}{1+\alpha+d}}$ and thus the lower bound (8). Let us note that this optimization is the same as in the Mott argument discussed above and that $E_c \downarrow 0$ and $r_c \uparrow \infty$ as $\beta \uparrow \infty$. Moreover, our optimized lower bound results from a critical resistor network roughly approximating the one appearing in [AHL].

The paper is organized as follows. In Sect. 2 we recall some definitions and results about point processes and state some technical results needed later on. The statements (i) and (ii) of Theorem 1 are proven in Sect. 3 and in Sects. 4 to 6, respectively. In Appendix A we show that the continuous-time random walk in the random environment is well defined by verifying the absence of explosion phenomena. Appendix B contains some technical proofs about the Palm measure. Appendix C is devoted to the proof of Proposition 1.

2. The Random Environment

In this section, we recall some properties of point processes (for more details, see [DV, FKAS, MKM, Kal, Tho]). In the sequel, given a topological space X, $\mathcal{B}(X)$ will denote the σ-algebra of Borel subsets of X. Given a set A, $|A|$ will denote its cardinality. Moreover, given a probability measure μ, we write $\mathbf{E}_\mu$ for the corresponding expectation.

2.1. Stationary simple marked point processes. Given a bounded complete separable metric space K, consider the space $\mathcal{N} := \mathcal{N}(\mathbb{R}^d \times K)$ of all counting measures ξ on $\mathbb{R}^d \times K$, *i.e.* integer-valued measures such that $\xi(B \times K) < \infty$ for any bounded set $B \in \mathcal{B}(\mathbb{R}^d)$. One can show that $\xi \in \mathcal{N}$ if and only if $\xi = \sum_j \delta_{(x_j,k_j)}$ where δ is the Dirac measure and $\{(x_j, k_j)\}$ is a countable family of (not necessarily distinct) points in $\mathbb{R}^d \times K$ with at most finitely many points in any bounded set. Then k_j is called the *mark* at x_j. Given $\xi \in \mathcal{N}$, we write $\hat{\xi} \in \mathcal{N}(\mathbb{R}^d)$ for the counting measure on $\mathbb{R}^d$ defined by $\hat{\xi}(B) = \xi(B \times K)$ for any $B \in \mathcal{B}(\mathbb{R}^d)$. Given $x \in \mathbb{R}^d$, we write $x \in \hat{\xi}$ whenever $x \in \mathrm{supp}(\hat{\xi})$. If $\hat{\xi}(\{x\}) \le 1$ for any $x \in \mathbb{R}^d$, we say that $\xi \in \mathcal{N}$ is *simple* and write $k_{x_j} := k_j$ for any $x_j \in \hat{\xi}$.

A metric on $\mathcal{N}$ can be defined in the following way [MKM, Sect. 1.15]. Fix an element $k^* \in K$. Denote by $B_r(x,k)$ and B_r the open balls in $\mathbb{R}^d \times K$ of radius $r > 0$ centred on (x,k) and on $(0,k^*)$, respectively. Let $\xi = \sum_{i\in I} \delta_{(x_i,k_i)}$ and $\xi' = \sum_{j\in J} \delta_{(x'_j,k'_j)}$ be elements of $\mathcal{N}$, where I, J are countable sets. Then ξ and ξ' are close to each other if any point (x_i, k_i) contained in B_n is close to a point (x'_j, k'_j) for arbitrary large n, up

to "boundary effects". More precisely, given a positive integer n, let $d_n(\xi, \xi')$ be the infimum over all $\varepsilon > 0$ such that there is a one-to-one map f from a (possibly empty) subset D of I into a subset of J with the properties:

(i) $\mathrm{supp}(\xi) \cap B_{n-\varepsilon} \subset \{(x_i, k_i) : i \in D\}$;
(ii) $\mathrm{supp}(\xi') \cap B_{n-\varepsilon} \subset \{(x'_j, k'_j) : j \in f(D)\}$;
(iii) $(x'_{f(i)}, k'_{f(i)}) \in B_\varepsilon(x_i, k_i)$ for $i \in D$.

One can show that $d_{\mathcal{N}}(\xi, \xi') = \sum_{n=1}^{\infty} 2^{-n} d_n(\xi, \xi')$ is a bounded metric on $\mathcal{N}$ and for this metric $\mathcal{N}$ is complete and separable. Moreover, the sets $\{\xi \in \mathcal{N} : \xi(B) = n\}$, $B \in \mathcal{B}(\mathbb{R}^d \times K)$, $n \in \mathbb{N}$, generate the Borel σ-algebra $\mathcal{B}(\mathcal{N})$ and $d_{\mathcal{N}}$ generates the coarsest topology such that $\xi \in \mathcal{N} \mapsto \int \xi(dx, dk)\, f(x, k)$ is continuous for any continuous function $f \geq 0$ on $\mathbb{R}^d \times K$ with bounded support. Finally, by choosing different reference points k^* one obtains equivalent metrics.

A *marked point process* on $\mathbb{R}^d$ with marks in K is then a measurable map Φ from a probability space into $\mathcal{N}$. We denote by $\mathcal{P}$ its distribution (a probability measure on $(\mathcal{N}, \mathcal{B}(\mathcal{N}))$). We say that the process is *simple* if $\mathcal{P}$-almost all $\xi \in \mathcal{N}$ are simple. The translations on $\mathbb{R}^d$ extend naturally to $\mathbb{R}^d \times K$ by $S_x : (y, k) \mapsto (x + y, k)$. This induces an action S of the translation group $\mathbb{R}^d$ on $\mathcal{N}$ by $(S_x \xi)(B) = \xi(S_x B)$, where $B \in \mathcal{B}(\mathbb{R}^d \times K)$ and $x \in \mathbb{R}^d$. For simple counting measures,

$$S_x \xi = \sum_{y \in \hat{\xi}} \delta_{(y-x, k_y)} \,.$$

A marked point process is said to be *stationary* if $\mathcal{P}(A) = \mathcal{P}(S_x A)$ for all $x \in \mathbb{R}^d$, $A \in \mathcal{B}(\mathcal{N})$, and *(space) ergodic* if the σ-algebra of translation invariant sets is trivial, *i.e.*, if $A \in \mathcal{B}(\mathcal{N})$ satisfies $S_x A = A$ for all $x \in \mathbb{R}^d$ then $\mathcal{P}(A) \in \{0, 1\}$. Due to [DV, Prop. 10.1.IV], if $\mathcal{P}$ is stationary and gives no weight to the trivial measure without any point (which will be assumed here), then

$$\mathcal{P}\big(\xi \in \mathcal{N} : \big|\,\mathrm{supp}(\hat{\xi})\,\big| = \infty\big) = 1 \,. \tag{10}$$

The marked point processes studied in this work are obtained by the procedure of randomization, which we recall now together with the related notion of thinning (see [Kal]). Let $\hat{\Phi}$ be a stationary simple point process (SSPP) on $\mathbb{R}^d$, ν be a probability measure on $[-1, 1]$ and $p \in [0, 1]$. The *ν–randomization* of $\hat{\Phi}$ is the stationary simple marked point process (SSMPP) Φ_ν obtained by assigning to each realization $\hat{\xi} = \sum_{i \in I} \delta_{x_i}$ of $\hat{\Phi}$ the measure $\xi = \sum_{i \in I} \delta_{(x_i, E_i)}$, where $\{E_i\}_{i \in I}$ are independent identically distributed random variables having distribution ν. Finally, the *p–thinning* $\hat{\Phi}_p$ of $\hat{\Phi}$ is the SSPP on $\mathbb{R}^d$ obtained by assigning to each realization $\hat{\xi}$ the measure $\sum_{i \in I} P_i\, \delta_{x_i}$, where $\{P_i\}_{i \in I}$ are independent Bernoulli variables with $\mathrm{Prob}(P_i = 1) = p$ and $\mathrm{Prob}(P_i = 0) = 1 - p$. Both the point processes Φ_ν and $\hat{\Phi}_p$ are examples of *stationary cluster processes*, also called *homogeneous cluster fields* (see [DV, Chap. 8] and [MKM, Chap. 10]). In particular, ergodicity is conserved by ν–randomization and p–thinning ([DV, Prop. 10.3.IX] and [MKM, Prop. 11.1.4]). To conclude, let us give a few examples.

Example 1. A *Poisson point process* (PPP) appears, as already discussed, naturally as limit distribution of thinnings. Given a measure μ on X, with X equal to $\mathbb{R}^d$ or $\mathbb{R}^d \times [-1, 1]$, the PPP on X with intensity measure μ is defined by the two conditions (i) for any $B \in \mathcal{B}(X)$, $\xi(B)$ is a Poisson random variable with expectation $\mu(B)$; (ii) for any

disjoint sets $B_1, \dots, B_n \in \mathcal{B}(X)$, $\xi(B_1), \dots, \xi(B_n)$ are independent. A PPP on $\mathbb{R}^d$ is stationary if and only if its intensity measure μ is proportional to the Lebesgue measure, $\mu = \rho\, dx$. In such a case it is an ergodic process satisfying the hypothesis (H2) of Theorem 1 and all moments ρ_κ, $\kappa > 0$ in (1) are finite. Its p–thinning is the PPP on $\mathbb{R}^d$ with intensity $p\rho$ while its ν–randomization is the PPP on $\mathbb{R}^d \times [-1, 1]$ with intensity measure $\rho\, dx \otimes \nu$.

Example 2. Let us associate to the uniformly distributed random variable y in the unit cube C_1 the point measure $\hat{\xi} = \sum_{z\in\mathbb{Z}^d} \delta_{z+y}$. The corresponding point process is an ergodic SSPP satisfying $\rho_\kappa = 1$ for any $\kappa > 0$. Although this process satisfies (H1), the SSMPP obtained from it via p-thinning and ν-randomization does not and does also not satisfy (H2). However, Theorem 1(ii) is still valid for this SSMPP, as can be checked by restricting the analysis of Sect. 6 to regions which are unions of boxes of the form $z + [0, 1)^d$, $z \in \mathbb{Z}^d$ and using the independence of $\hat{\xi}(A)$ and $\hat{\xi}(B)$ when A, B are disjoint unions of such boxes.

Other examples of ergodic SSMPP can be obtained by means of SSPP with short–range correlations (see [DV, Exercise 10.3.4]). Of particular relevance for solid state physics are point processes associated to random or quasiperiodic tilings [BHZ], which satisfy the hypothesis (H1) of Theorem 1.

2.2. The Palm distribution. In what follows, it will always be assumed that $\hat{\mathcal{P}}$ and $\mathcal{P}$ are defined as in Theorem 1 and that (6) is satisfied if ν is a Dirac measure. In order to shorten notations, we will write $\mathcal{N}$ and $\hat{\mathcal{N}}$ for $\mathcal{N}(\mathbb{R}^d \times [-1, 1])$ and $\mathcal{N}(\mathbb{R}^d)$, respectively. We would like now to "pick up at random" a point among $\{x_j\}$ and take it as the origin. One thus looks at the following borelian subset of $\mathcal{N}$:

$$\mathcal{N}_0 := \left\{ \xi \in \mathcal{N} : 0 \in \hat{\xi} \right\} .$$

Since $\mathcal{N}_0$ is closed, it defines a bounded complete separable metric space. Note that $x \in \hat{\xi}$ if and only if $S_x\xi \in \mathcal{N}_0$. The Palm distribution $\mathcal{P}_0$ on $\mathcal{N}_0$ associated to $\mathcal{P}$ is now defined as follows. Consider the measurable map $\mathcal{G}$ from $\mathcal{N}$ into $\mathcal{N}(\mathbb{R}^d \times \mathcal{N}_0)$ given by $\xi \mapsto \xi^* = \sum_{x\in\hat{\xi}} \delta_{(x,S_x\xi)}$. Let $\mathcal{P}^* = \mathcal{G}_*\mathcal{P}$ be the distribution of the marked point process on $\mathbb{R}^d \times \mathcal{N}_0$ with mark space $\mathcal{N}_0$, namely $\mathcal{P}^*$ is the image under $\mathcal{G}$ of the probability measure $\mathcal{P}$ on $\mathcal{N}$. It is easy to show that $\mathcal{G} \circ S_x = S_x^* \circ \mathcal{G}$ for $x \in \mathbb{R}^d$, where S_x^* is the action on $\mathbb{R}^d \times \mathcal{N}_0$ of the translations given by $(y, \xi) \mapsto (y + x, \xi)$. As a result, $\mathcal{P}^*$ is also stationary. Then, for any fixed $A \in \mathcal{B}(\mathcal{N}_0)$, the measure $\mu_A(B) = \int \mathcal{P}^*(d\xi^*)\, \xi^*(B \times A)$ on $\mathbb{R}^d$ is translation invariant and thus proportional to the Lebesgue measure. This implies that, for any $N > 0$ and any $A \in \mathcal{B}(\mathcal{N}_0)$,

$$\mathcal{C}_\mathcal{P}(A) := \int_{\mathcal{N}(\mathbb{R}^d\times\mathcal{N}_0)} \mathcal{P}^*(d\xi^*)\, \xi^*(C_1 \times A) = \frac{1}{N^d} \int_{\mathcal{N}(\mathbb{R}^d\times\mathcal{N}_0)} \mathcal{P}^*(d\xi^*)\, \xi^*(C_N \times A).$$

The *Palm distribution* associated to $\mathcal{P}$ is the probability measure $\mathcal{P}_0$ on $\mathcal{N}_0$ obtained from $\mathcal{C}_\mathcal{P}$ by normalization, namely, $\mathcal{P}_0 = \rho^{-1}\mathcal{C}_\mathcal{P}$, where ρ is defined in (1). Thus, for any $N > 0$,

$$\mathcal{P}_0(A) := \frac{1}{\rho}\frac{1}{N^d} \int_\mathcal{N} \mathcal{P}(d\xi) \int_{C_N} \hat{\xi}(dx)\, \chi_A(S_x\xi) , \tag{11}$$

where χ_A is the characteristic function on the Borel set $A \subset \mathcal{N}_0$. One can show [FKAS, Theorem 1.2.8] that for any nonnegative measurable function f on $\mathbb{R}^d \times \mathcal{N}_0$,

$$\int_{\mathbb{R}^d} dx \int_{\mathcal{N}_0} \mathcal{P}_0(d\xi)\ f(x,\xi) \ = \ \frac{1}{\rho} \int_{\mathcal{N}} \mathcal{P}(d\xi) \int_{\mathbb{R}^d} \hat{\xi}(dx)\ f(x, S_x\xi)\,, \tag{12}$$

which is used in [DV] as the definition of $\mathcal{P}_0$. Similarly, there is a Palm distribution $\hat{\mathcal{P}}_0$ on $\hat{\mathcal{N}}_0 := \{\hat{\xi} \in \hat{\mathcal{N}} : 0 \in \hat{\xi}\}$ associated to the distribution $\hat{\mathcal{P}}$ of a SSPP on $\mathbb{R}^d$.

It is known that the Palm distribution of a stationary PPP on $\mathbb{R}^d$ with distribution $\hat{\mathcal{P}}$ (Example 1 above) is the convolution $\hat{\mathcal{P}}_0 = \hat{\mathcal{P}} * \delta_{\delta_0}$ of $\hat{\mathcal{P}}$ with the Dirac measure at $\hat{\xi} = \delta_0$ (*i.e.* $\hat{\mathcal{P}}_0$ is simply obtained by adding a point at the origin). The Palm distribution of a PPP on $\mathbb{R}^d \times [-1, 1]$ with intensity measure $\rho\, dx \otimes \nu$ is the convolution $\mathcal{P}_0 = \mathcal{P} * \zeta$, where ζ is the distribution of a marked point process obtained by ν–randomization of δ_{δ_0}. The Palm distribution associated to the SSPP in Example 2 is $\hat{\mathcal{P}}_0 = \delta_{\sum_{x\in\mathbb{Z}^d} \delta_x}$. Its ν–randomization is the Palm distribution of the ν–randomization of Example 2.

We collect in the lemma below a number of results on the Palm distribution which will be needed in the sequel. Their proofs are given in Appendix B.

Lemma 1. *(i) Let $k : \mathcal{N}_0 \times \mathcal{N}_0 \to \mathbb{R}$ be a measurable function such that $\int \hat{\xi}(dx)\, |k(\xi, S_x\xi)|$ and $\int \hat{\xi}(dx)\, |k(S_x\xi, \xi)|$ are in $L^1(\mathcal{N}_0, \mathcal{P}_0)$. Then*

$$\int \mathcal{P}_0(d\xi) \int \hat{\xi}(dx)\ k(\xi, S_x\xi) \ = \ \int \mathcal{P}_0(d\xi) \int \hat{\xi}(dx)\ k(S_x\xi, \xi)\,.$$

(ii) Let $\Gamma \in \mathcal{B}(\mathcal{N})$ be such that $S_x\Gamma = \Gamma$ for all $x \in \mathbb{R}^d$. Then $\mathcal{P}(\Gamma) = 1$ if and only if $\mathcal{P}_0(\Gamma_0) = 1$ with $\Gamma_0 = \Gamma \cap \mathcal{N}_0$.

(iii) Let $\mathcal{P}$ be ergodic and $A, B \in \mathcal{B}(\mathcal{N}_0)$ be such that $B \subset A$, $\mathcal{P}_0(A \setminus B) = 0$ and $S_x\xi \in A$ for any $\xi \in B$ and any $x \in \hat{\xi}$. Then $\mathcal{P}_0(A) \in \{0, 1\}$.

(iv) Let $A_j \in \mathcal{B}(\mathbb{R}^d)$ for $j = 1, \ldots, n$. Then

$$\mathbf{E}_{\mathcal{P}_0}\left(\prod_{j=1}^{n} \hat{\xi}(A_j)\right) \ \le \ \frac{c}{\rho}\, \mathbf{E}_{\mathcal{P}}\big(\hat{\xi}(C_1)^{n+1}\big) \ + \ \frac{c}{\rho} \sum_{j=1}^{n} \mathbf{E}_{\mathcal{P}}\big(\hat{\xi}(\tilde{A}_j)^{n+1}\big)\,, \tag{13}$$

where $\tilde{A}_j := \cup_{x\in C_1}\big(A_j + x\big)$ and c is a positive constant depending on n.

Remark 1. Here we point out a simple geometric property of point measures ξ within the set

$$\mathcal{W} := \big\{\xi \in \mathcal{N}_0 \,:\, S_x\xi \neq \xi\ \forall x \in \mathbb{R}^d \setminus \{0\}\big\}, \tag{14}$$

which will be fundamental in order to apply the methods developed in [KV] and [DFGW]. Let us consider a sequence $\{x_n\}_{n\geq 0}$ of elements in $\mathrm{supp}(\hat{\xi})$ with $x_0 = 0$ and set $\xi_n := S_{x_n}\xi$. The ξ_n can be thought of as the *environment viewed from the point x_n*. Due to the definition of $\mathcal{W}$, $\{x_n\}_{n\geq 0}$ can be recovered from $\{\xi_n\}_{n\in\mathbb{N}}$ by means of the identities

$$x_{n+1} - x_n \ = \ \Delta(\xi_n, \xi_{n+1}), \qquad n \in \mathbb{N}\,,$$

where the function $\Delta : \mathcal{W} \times \mathcal{N}_0 \to \mathbb{R}^d$ is defined as

$$\Delta(\xi', \xi'') \ := \ \begin{cases} x & \text{if } \xi'' = S_x\xi'\,, \\ 0 & \text{otherwise}\ . \end{cases} \tag{15}$$

Note that, by Lemma 1(ii), condition (6) is equivalent to $\mathcal{P}_0(\mathcal{W}) = 1$.

3. Variational Formula

The main object of this section is to show the following result, implying Theorem 1(i).

Theorem 2. *Let $\mathcal{P}$ satisfy the assumptions of* Theorem 1(i)*. Then the limit* (3) *exists and D is given by the variational formula*

$$(a \cdot D\, a) = \inf_{f \in L^\infty(\mathcal{N}_0, \mathcal{P}_0)} \int \mathcal{P}_0(d\xi) \int \hat{\xi}(dx) c_{0,x}(\xi) \big(a \cdot x + \nabla_x f(\xi)\big)^2, \quad a \in \mathbb{R}^d, \tag{16}$$

with

$$\nabla_x f(\xi) \;:=\; f(S_x \xi) - f(\xi)\,. \tag{17}$$

Moreover, the rescaled process $\underline{Y}^{\xi,\varepsilon} := (\varepsilon X^{\xi}_{t\varepsilon^{-2}})_{t\geq 0}$ defined on $(\Omega_\xi, \mathbf{P}^\xi_0)$ converges weakly in $\mathcal{P}_0$–probability as $\varepsilon \to 0$ to a Brownian motion $\underline{W}_D$ with covariance matrix D.

The proof is based on the theory of Ref. [KV] and [DFGW] and, in particular, Theorem 2.2 of [DFGW]. Because of the geometric disorder and the possibility of jumps between any of the random points, the application of this general theorem to our model is technically considerably more involved than in the case of the lattice model with jumps to nearest neighbors studied in [DFGW, Sect. 4]. As a preamble, let us state a result on the process X^ξ_t proven in Appendix C which will be used several times below.

Proposition 1. *Let $\mathcal{P}$ satisfy $\rho_\kappa < \infty$ for some integer $\kappa > 3$. Then, given $t > 0$ and $0 < \gamma < \kappa - 3$,*

$$\mathbf{E}_{\mathcal{P}_0} \mathbf{E}_{\mathbf{P}^\xi_0} \left(|X^\xi_t|^\gamma \right) \;<\; \infty\,.$$

Remark 2. From the variational formula of the diffusion matrix D given in Theorem 2 one can easily prove (see *e.g.* [DFGW]) that D is a multiple of the identity whenever $\mathcal{P}$ is isotropic (*i.e.*, it is invariant under all rotations by $\pi/2$ in a coordinate plane). In this case, the arguments leading to a lower bound on D are slightly simpler (and can be easily adapted to the general case). Therefore, in order to simplify the discussion and without loss of generality, in the last Sects. 5 and 6 we will assume $\mathcal{P}$ to be isotropic.

3.1. The result of De Masi, Ferrari, Goldstein and Wick. A main idea in [DFGW] is to study the process $(S_{X^\xi_t}\xi)_{t\geq 0}$ with values in the space $\mathcal{N}_0$ of the environment configurations, instead of the random walk $(X^\xi_t)_{t\geq 0}$. This process is called the process *environment viewed from the particle*. It is defined on the probability space $(\Omega_\xi, \mathbf{P}^\xi_0)$, with $\Omega_\xi = D([0,\infty), \mathrm{supp}(\hat{\xi}))$. Let $\mathbf{P}_\xi$ be its distribution on the path space $\Xi := D([0,\infty), \mathcal{N}_0)$ (endowed as usual with the Skorohod topology). A generic element of Ξ will be denoted by $\underline{\xi} = (\xi_t)_{t\geq 0}$. Let us set $\mathbf{P} := \int \mathcal{P}_0(d\xi)\mathbf{P}_\xi$. The *environment process* is the process $(\xi_t)_{t\geq 0}$ defined on the probability space $(\Xi, \mathbf{P})$ with distribution $\mathbf{P}$. This is a continuous–time jump Markov process with initial measure $\mathcal{P}_0$ and transition probabilities

$$\mathbf{P}(\xi_{s+t} = \xi' \,|\, \xi_s = \xi) \;=\; \mathbf{P}_\xi(\xi_t = \xi') \;=:\; p_t(\xi'|\xi) \qquad \forall\, s, t \geq 0$$

with, for any $\xi \in \mathcal{W}$,

$$p_t(\xi'|\xi) \;=\; \begin{cases} p_t^{\xi}(x|0) & \text{if } \xi' = S_x\xi \text{ for some } x \in \hat{\xi}\ , \\ 0 & \text{otherwise}\ . \end{cases} \tag{18}$$

For any time $t \geq 0$, let us introduce the random variable $X_t : \Xi \to \mathbb{R}^d$ defined by

$$X_t(\underline{\xi}) \;:=\; \sum_{s\in[0,t]} \Delta_s(\underline{\xi})\ , \tag{19}$$

where

$$\Delta_s(\underline{\xi}) \;:=\; \begin{cases} x & \text{if } \xi_s \;=\; S_x\xi_{s^-} \\ 0 & \text{otherwise} \end{cases}$$

and the sum runs over all jump times s for which $\Delta_s(\underline{\xi}) \neq 0$. Note that $\{X_{[s,t]} := X_t - X_s : t > s \geq 0\}$ defines an antisymmetric additive covariant family of random variables as defined in [DFGW], and X_t has paths in $D([0,\infty), \mathbb{R}^d)$. The crucial link to the dynamics of a particle in a fixed environment is now the following: due to Remark 1, for any $\xi \in \mathcal{W}$, the distribution of the process $(X_t)_{t\geq 0}$ defined on $(\Xi, \mathbf{P}_\xi)$ is equal to the distribution $\mathbf{P}_0^{\xi}$ of the random walk on $\mathrm{supp}(\hat{\xi})$ (naturally embedded in $\mathbb{R}^d$) starting at the origin. Recalling that $\mathbf{P} = \int \mathcal{P}_0(d\xi)\, \mathbf{P}_\xi$, this implies

$$\mathbf{E}_{\mathcal{P}_0}\left(\mathbf{E}_{\mathbf{P}_0^{\xi}}\left((X_t^{\xi}\cdot a)^2\right)\right) \;=\; \mathbf{E}_{\mathbf{P}}\left((X_t\cdot a)^2\right)\ , \tag{20}$$

which gives a way to calculate the diffusion matrix D from the distribution $\mathbf{P}$ on Ξ.

In order to apply Theorem 2.2 of [DFGW], it is enough to verify the following hypothesis:

(a) the environment process is reversible and ergodic;
(b) the random variables $X_{[s,t]}$, $0 \leq s < t$ are in $L^1(\Xi, \mathbf{P})$;
(c) the mean forward velocity exists:

$$\varphi(\xi) \;:=\; L^2\text{–}\lim_{t\downarrow 0}\ \frac{1}{t}\ \mathbf{E}_{\mathbf{P}_\xi}(X_t)\ . \tag{21}$$

(d) the martingale $X_t - \int_0^t ds\, \varphi(\xi_s)$ is in $L^2(\Xi, \mathbf{P})$.

Let us assume $\rho_{12} < \infty$. Then, statement (a) will be proved in Proposition 2, Subsect. 3.3. The statement (b) follows from Proposition 1. The L^2-convergence in (c) will be proved in Subsect. 3.4 (Proposition 4), where we also show the L^2-convergence in the following formula defining the mean square displacement matrix $\psi(\xi)$:

$$(a\cdot\psi(\xi)a) \;:=\; L^2\text{–}\lim_{t\downarrow 0}\ \frac{1}{t}\ \mathbf{E}_{\mathbf{P}_\xi}\left((a\cdot X_t)^2\right)\ . \tag{22}$$

The last point (d) is a consequence of Proposition 1 assuring that $X_t \in L^2(\Xi, \mathbf{P})$ and the fact that $\int_0^t ds\, \varphi(\xi_s) \in L^2(\mathcal{N}_0, \mathcal{P}_0)$, which can be proved by means of the Cauchy–Schwarz inequality, the stationarity of $\mathbf{P}$ following from (a), and the property $\varphi \in L^2(\mathcal{N}, \mathcal{P}_0)$.

Once hypotheses (a)-(d) have been verified, one can invoke [DFGW, Theorem 2.2 and Remark 4, p. 802] to conclude that limit (3) exists and that the rescaled random walk $\underline{Y}^{\xi,\varepsilon}$ converges weakly in $\mathcal{P}_0$–probability to the Brownian motion $\underline{W}_D$ with covariance matrix D given by (3), and that D is moreover given by

$$(a \cdot Da) = \mathbf{E}_{\mathcal{P}_0}\big((a \cdot \psi a) \big) - 2 \int_0^\infty dt \, \Big\langle \varphi \cdot a \, , \, e^{t\mathcal{L}} \, \varphi \cdot a \Big\rangle_{\mathcal{P}_0} \, , \tag{23}$$

where $\mathcal{L}$ is the generator of the environment process and the integral on the r.h.s. is finite. Formula (16) can be deduced from the expressions of $\mathcal{L}$, φ and ψ established in the following subsections (Propositions 3 and 4) by using a known general result on self-adjoint operators stated in (47) below.

3.2. Preliminaries. Before starting to prove the above-mentioned statements (a)-(d), let us fix some notations and recall some general facts about jump Markov processes. In what follows, given a complete separable metric space Z we denote by $\mathcal{F}(Z)$ the family of bounded Borel functions on Z and, given a (not necessarily finite) interval $I \subset \mathbb{R}$, we denote by $D(I, Z)$ the space of right continuous paths $\underline{z} = (z_t)_{t\in I}$, $z_t \in Z$, having left limits. The path space $D(I, Z)$ is endowed with the Skorohod topology [Bil] which is the natural choice for the study of jump Markov processes. For a time $s \geq 0$, the time translation τ_s is defined as

$$\tau_s : D([0,\infty), Z) \to D([0,\infty), Z), \qquad (\tau_s \underline{z})_t \; := \; z_{t+s} \, .$$

Moreover, given $0 \leq a < b$, we denote by $R_{[a,b]}$ the function

$$R_{[a,b]} \; : \; D([0,\infty), Z) \to D([a,b], Z), \qquad (R_{[a,b]}\underline{z})_t \; := \; \lim_{\delta \downarrow 0} \; z_{a+b-t-\delta} \, .$$

$R_{[a,b]}\underline{z}$ is the time–reflection of $(z_t)_{t\in[a,b]}$ w.r.t. the middle point of $[a, b]$, and it can naturally be extended to paths on $[0, a + b]$.

A continuous–time Markov process with path in $D([0,\infty), Z)$ and distribution $\mathbf{p}$ is called *stationary* if $\mathbf{E}_\mathbf{p}(F) = \mathbf{E}_\mathbf{p}(F \circ \tau_s)$ for all $s \geq 0$ and for any bounded Borel function F on $D([0,\infty), Z)$. It is called *reversible* if $\mathbf{E}_\mathbf{p}(F) = \mathbf{E}_\mathbf{p}(F \circ R_{[a,b]})$ for all $b > a \geq 0$ and any bounded Borel function F on $D([0,\infty), Z)$ such that $F(\underline{z})$ depends only on $(z_t)_{t\in[a,b]}$. Thanks to the Markov property, one can show that stationarity is equivalent to

$$\mathbf{E}_\mathbf{p}\big(f(z_0)\big) = \; \mathbf{E}_\mathbf{p}\big(f(z_s)\big) \, , \qquad \forall \; s \geq 0 \, , \;\; \forall \; f \in \mathcal{F}(Z), \tag{24}$$

while reversibility is equivalent to

$$\mathbf{E}_\mathbf{p}\big(f(z_0)g(z_s)\big) = \; \mathbf{E}_\mathbf{p}\big(g(z_0)f(z_s)\big) \, , \qquad \forall \; s \geq 0 \, , \;\; \forall \; f, g \in \mathcal{F}(Z) \, . \tag{25}$$

In particular, stationarity follows from reversibility. Finally, the Markov process is called *(time) ergodic* if $\mathbf{p}(A) \in \{0, 1\}$ whenever $A \in \mathcal{B}\big(D([0,\infty), Z)\big)$ is time-shift invariant, i.e. $A = \tau_s A$ for all $s \geq 0$. Recall that if the Markov process is stationary then it can be extended to a Markov process with path space $D(\mathbb{R}, Z)$ and the resulting distribution is univocally determined (this follows from Kolmogorov's extension theorem and the regularity of paths). Now stationarity, reversibility and ergodicity of the extended process are defined as above by means of τ_s, $s \in \mathbb{R}$, and $R_{[a,b]}$, $-\infty < a < b < \infty$. Then one can check that these properties are preserved by extension (for what concerns ergodicity,

see in particular [Ros, Chapter 15, p. 96–97]). Therefore our definitions coincide with those in [DFGW].

All the above definitions and remarks can be extended in a natural way to discrete–time Markov processes (with path space $Z^{\mathbb{N}}$). Moreover, in the discrete case, stationarity and reversibility are equivalent respectively to (24) and (25) with $s = 1$.

We conclude this section recalling the standard construction of the continuous–time random walk satisfying conditions (C1) and (C2) in the Introduction. We first note that these conditions are meaningful for $\mathcal{P}_0$–almost all ξ if $\mathbf{E}_{\mathcal{P}_0}(\lambda_0) < \infty$. In fact, due to the bound $\lambda_x(\xi) \le e^{4\beta}\, e^{|x|}\, \lambda_0(\xi)$, one can infer from $\mathbf{E}_{\mathcal{P}_0}(\lambda_0) < \infty$ that $\lambda_x(\xi) < \infty$ for any $x \in \hat{\xi}$, $\mathcal{P}_0$ a.s. We note that the condition $\mathbf{E}_{\mathcal{P}_0}(\lambda_0) < \infty$ is equivalent to $\rho_2 < \infty$ due to the following lemma:

Lemma 2. *For any positive integer k, $\mathbf{E}_{\mathcal{P}_0}(\lambda_0^k) < \infty$ if and only if $\rho_{k+1} < \infty$.*

Proof. Note that for suitable positive constants c_1, c_2 one has

$$c_1 \sum_{z\in\mathbb{Z}^d} \hat{\xi}(C_1+z)e^{-|z|} \;\le\; \lambda_0(\xi) \;\le c_2 \sum_{z\in\mathbb{Z}^d} \hat{\xi}(C_1+z)e^{-|z|}\,, \qquad \mathcal{P}_0\text{-a.s.}$$

Next let us expand the k^{th} power of these inequalities. By applying Lemma 1(iv) and using the stationarity of $\mathcal{P}$, one gets that $\mathbf{E}_{\mathcal{P}_0}(\lambda_0^k) < \infty$ if $\rho_{k+1} < \infty$. Suppose now that $\mathbf{E}_{\mathcal{P}_0}(\lambda_0^k) < \infty$. Then the above expansion in k^{th} power implies that $\mathbf{E}_{\mathcal{P}_0}(\hat{\xi}(C_1)^k) < \infty$. Since due to (11),

$$\mathbf{E}_{\mathcal{P}_0}(\hat{\xi}(C_1)^k) \;=\; \frac{2^d}{\rho}\int_{\mathcal{N}} \mathcal{P}(d\xi) \int_{C_{1/2}} \hat{\xi}(dx)\, \hat{\xi}(C_1+x)^k \ge \frac{2^d}{\rho}\, \mathbf{E}_{\mathcal{P}}(\hat{\xi}(C_{1/2})^{k+1})\,,$$

one concludes that $\mathbf{E}_{\mathcal{P}}(\hat{\xi}(C_{1/2})^{k+1}) < \infty$, which is equivalent to $\rho_{k+1} < \infty$. □

The construction of the continuous–time random walk follows standard references (*e.g.* [Bre, Chap. 15] and [Kal, Chap. 12]) and can be described roughly as follows: After arriving at site $y \in \hat{\xi}$, the particle waits an exponential time with parameter $\lambda_y(\xi)$ and then jumps to another site $z \in \hat{\xi}$ with probability

$$p^{\xi}(z|y) \;:=\; \frac{c_{y,z}(\xi)}{\lambda_y(\xi)}\,. \tag{26}$$

More precisely, consider $\xi \in \mathcal{N}_0$ such that $0 < \lambda_z(\xi) < \infty$ for any $z \in \hat{\xi}$ and set $\tilde{\Omega}_\xi := \big(\mathrm{supp}(\hat{\xi})\big)^{\mathbb{N}}$. A generic path in $\tilde{\Omega}_\xi$ is denoted by $\big(\tilde{X}_n^\xi\big)_{n\ge 0}$. Given $x \in \hat{\xi}$, let $\tilde{\mathbf{P}}_x^\xi$ be the distribution on $\tilde{\Omega}_\xi$ of a discrete–time random walk on $\mathrm{supp}(\hat{\xi})$ starting in x and having transition probabilities $p^\xi(z|y)$. Let $\big(\Theta, \mathbf{Q}\big)$ be another probability space where the random variables $T_{n,z}^\xi$, $z \in \hat{\xi}$, $n \in \mathbb{N}$, are independent and exponentially distributed with parameter $\lambda_z(\xi)$, namely $\mathbf{Q}\big(\,T_{z,n}^\xi > t\,\big) = \exp\big(-\lambda_z(\xi)t\big)$. On the probability space $(\tilde{\Omega}_\xi \times \Theta, \tilde{\mathbf{P}}_x^\xi \otimes \mathbf{Q})$ define the following functions:

$$R_0^\xi \;:=\; 0\,; \quad R_n^\xi := T^\xi_{0,\tilde{X}_0^\xi} + T^\xi_{1,\tilde{X}_1^\xi} + \cdots + T^\xi_{n-1,\tilde{X}_{n-1}^\xi} \quad \text{if} \quad n \ge 1\,,$$
$$n_*^\xi(t) \;:=\; n \quad \text{if} \quad R_n^\xi \le t < R_{n+1}^\xi\,.$$

Note that $n_*^\xi(t)$ is well posed for any $t \geq 0$ only if $\lim_{n\uparrow\infty} R_n^\xi = \infty$. If

$$\tilde{\mathbf{P}}_x^\xi \otimes \mathbf{Q}\Big(\lim_{n\uparrow\infty} R_n^\xi = \infty \Big) \;=\; 1 \;, \tag{27}$$

then, due to [Bre, Theorem 15.37], the random walk $(\tilde{X}^\xi_{n_*^\xi(t)})_{t\geq 0}$, defined $\tilde{\mathbf{P}}_x^\xi \otimes \mathbf{Q}$–almost everywhere, is a jump Markov process whose distribution satisfies the infinitesimal conditions (C1) and (C2). The condition $\lim_{n\uparrow\infty} R_n^\xi = \infty$ assures that no *explosion phenomenon* takes place, notably only finitely many jumps can occur in finite time intervals. In Appendix A we prove that (27) is verified if $\rho_2 < \infty$.

3.3. The environment viewed from the particle. The process environment viewed from the particle and the environment process have been introduced in Sect. 3.1. Given $t > 0$, we write $n_*(t)$ for the function on the path space $\Xi = D([0,\infty), \mathcal{N}_0)$ associating to each $\xi \in \Xi$ the corresponding number of jumps in the time interval $[0, t]$. Motivated by further applications, it is convenient to consider also the discrete–time versions of the above processes. Consider the discrete-time Markov process $\big(S_{\tilde{X}_n^\xi}\xi \big)_{n\geq 0}$ defined on $\big(\tilde{\Omega}_\xi, \tilde{\mathbf{P}}_0^\xi \big)$, call $\tilde{\mathbf{P}}_\xi$ its distribution on the path space $\tilde{\Xi} := \mathcal{N}_0^{\mathbb{N}}$ and denote a generic path in $\tilde{\Xi}$ by $(\xi_n)_{n\geq 0}$. Such a Markov process can be thought of as the environment viewed from the particle performing the discrete–time random walk with distribution $\tilde{\mathbf{P}}_0^\xi$. Let us point out a few properties of the distribution $\tilde{\mathbf{P}}_\xi$. First, we remark that due to the covariant relations

$$c_{z,y}(S_x\xi) \;=\; c_{z+x,y+x}(\xi) \;, \qquad \lambda_y(S_x\xi) \;=\; \lambda_{y+x}(\xi), \tag{28}$$

the process $\big(\lambda_{\tilde{X}_n^\xi}(\xi)\big)_{n\in\mathbb{N}}$ defined on $(\tilde{\Omega}_\xi, \tilde{\mathbf{P}}_0^\xi)$ and the process $(\lambda_0(\xi_n))_{n\in\mathbb{N}}$ defined on $(\tilde{\Xi}, \tilde{\mathbf{P}}_\xi)$ have the same distribution. Moreover, due to Remark 1, if $\xi \in \mathcal{W}$, then the process $(\zeta_n)_{n\in\mathbb{N}}$ defined on $(\tilde{\Xi}, \tilde{\mathbf{P}}_\xi)$ by $\zeta_0 = 0$ and

$$\zeta_n \;=\; \sum_{k=0}^{n-1} \Delta(\xi_k, \xi_{k+1}) \;, \forall\, n \geq 1 \;,$$

where $\Delta(\xi, \xi')$ is given by (15), has paths in $\tilde{\Omega}_\xi$ with distribution $\tilde{\mathbf{P}}_0^\xi$. Finally, it is convenient to consider a suitable average of the distributions $\tilde{\mathbf{P}}_\xi$. To this aim, let $\mathcal{Q}_0$ be the probability measure on $\mathcal{N}_0$ defined as

$$\mathcal{Q}_0(d\xi) \;:=\; \frac{\lambda_0(\xi)}{\mathbf{E}_{\mathcal{P}_0}(\lambda_0)}\, \mathcal{P}_0(d\xi) \;,$$

and set $\tilde{\mathbf{P}} := \int \mathcal{Q}_0(d\xi)\tilde{\mathbf{P}}_\xi$. If $\xi \in \mathcal{W}$, the transition probabilities are

$$p(\xi'|\xi) \;:=\; \tilde{\mathbf{P}}\big(\xi_{n+1} = \xi' \,|\xi_n = \xi\big) \;=\; \begin{cases} \lambda_0^{-1}(\xi) c_{0,x}(\xi) & \text{if } \xi' = S_x\xi \;, \\ 0 & \text{otherwise} \;. \end{cases}$$

Note that, due to (28) and the symmetry of the jump rates (2), $\lambda_0(\xi)p(\xi'|\xi) = \lambda_0(\xi')p(\xi|\xi')$.

Proposition 2. *Let $\rho_2 < \infty$. Then the process $(\xi_t)_{t\geq 0}$ defined on $(\Xi, \mathbf{P})$ is reversible, i.e.*

$$\mathbf{E}_{\mathbf{P}}\big(f(\xi_0)g(\xi_t)\big) \;=\; \mathbf{E}_{\mathbf{P}}\big(g(\xi_0)f(\xi_t)\big) \qquad \forall\; f, g \in \mathcal{F}(\mathcal{N}_0)\;,\;\; \forall\, t > 0\;, \tag{29}$$

and is (time) ergodic if $\mathcal{P}$ is ergodic. Similarly, the discrete-time Markov process $(\xi_n)_{n\geq 0}$ defined on $(\tilde{\Xi}, \tilde{\mathbf{P}})$ is reversible and is (time) ergodic.

Having at our disposal Lemma 1, the proof follows modifying arguments of *e.g.* [DFGW].

Proof. We give the proof for the continuous–time process, the discrete–time case being similar. We first verify the symmetric property $p_t(\xi'|\xi) = p_t(\xi|\xi')$. Actually, thanks to the construction of the dynamics given in Sect. 3.2, one can show that for any positive integer n and any $\xi = \xi^{(0)}, \xi^{(1)}, \ldots, \xi^{(n-1)}, \xi^{(n)} = \xi' \in \mathcal{N}_0$,

$$\mathbf{P}_\xi\big(n_*(t) = n, \xi_{R_1}{=}\xi^{(1)}, \ldots, \xi_{R_n}{=}\xi^{(n)}\big){=}\mathbf{P}_{\xi'}\big(n_*(t){=}n, \xi_{R_1} = \xi^{(n-1)}, \ldots, \xi_{R_n}{=}\xi^{(0)}\big),$$

where, given $\underline{\xi} \in \Xi$, $R_1(\underline{\xi}) < R_2(\underline{\xi}) < \ldots$ denote the jump times of the path $\underline{\xi}$. Next, given $f, g \in \mathcal{F}(\mathcal{N}_0)$ one gets by applying Lemma 1(i) and using $p_t(\xi'|\xi) = p_t(\xi|\xi')$ that

$$\begin{aligned}&\int \mathcal{P}_0(d\xi)\int \hat{\xi}(dx)\; p_t(S_x\xi|\xi)\; f(\xi)g(S_x\xi)\\ &\quad= \int \mathcal{P}_0(d\xi)\int \hat{\xi}(dx)\; p_t(S_x\xi|\xi)\; f(S_x\xi)g(\xi),\end{aligned} \tag{30}$$

which is equivalent to (29). Hence $\mathbf{P}$ is reversible. Due to Corollary 5 in [Ros, Chap. IV], in order to prove ergodicity it is enough to show that $\mathcal{P}_0(A) \in \{0, 1\}$ if $A \in \mathcal{B}(\mathcal{N}_0)$ has the following property: $\mathbf{P}_\xi(\xi_t \in A) = \chi_A(\xi)$ for $\mathcal{P}_0$–almost all ξ. Given such a set A, then there exists a Borel subset $B \subset A$ such that $\mathcal{P}_0(A \setminus B) = 0$ and $\mathbf{P}_\xi(\xi_t \in A) = 1$ for any $\xi \in B$. Fix $\xi \in B$ and $x \in \hat{\xi}$, then $\mathbf{P}_\xi(\xi_t = S_x\xi, \xi_t \in A) = \mathbf{P}_\xi(\xi_t = S_x\xi) > 0$ (the last bound follows from the positivity of the jump rates). Hence $S_x\xi \in A$. Lemma 1(iii) implies that $\mathcal{P}_0(A) \in \{0, 1\}$, thus allowing to conclude the proof. □

Let $\mathcal{P}$ fulfill the assumption of Proposition 2. Then,

$$(\mathcal{T}_t f)(\xi) \;:=\; \mathbf{E}_{\mathbf{P}_\xi}\big(f(\xi_t)\big) \;=\; \int \hat{\xi}(dx)\; p_t(S_x\xi|\xi)\, f(S_x\xi)\;, \qquad \mathcal{P}_0 \text{ a.s.} \tag{31}$$

defines a strongly continuous contraction semigroup on $L^2(\mathcal{N}_0, \mathcal{P}_0)$ (Markov semigroup). Actually, (i) $\mathcal{T}_t : L^2(\mathcal{N}_0, \mathcal{P}_0) \to L^2(\mathcal{N}_0, \mathcal{P}_0)$ is self-adjoint by (29) and is a contraction by the Cauchy-Schwarz inequality and the stationarity of $\mathbf{P}$; (ii) $\mathcal{T}_{t+s} = \mathcal{T}_t\mathcal{T}_s$ follows from the Markov nature of the process; (iii) the continuity follows from the following argument: first observe that it is enough to prove the continuity of $\mathcal{T}_t f$ at $t = 0$ for $f \in L^\infty(\mathcal{N}_0, \mathcal{P}_0)$, which is obtained from the dominated convergence theorem and the estimate $|(\mathcal{T}_t f - f)(\xi)| \leq 2\|f\|_\infty(1 - p_t^\xi(0|0))$.

Let us denote by $\mathcal{L}$ the generator of the Markov semigroup $(\mathcal{T}_t)_{t\geq 0}$ and by $D(\mathcal{L}) \subset L^2(\mathcal{N}_0, \mathcal{P}_0)$ its domain.

Proposition 3. *Let $\mathcal{P}$ satisfy $\rho_4 < \infty$. Then $\mathcal{L}$ is nonpositive and self–adjoint with core $L^\infty(\mathcal{N}_0, \mathcal{P}_0)$. For any $f \in L^\infty(\mathcal{N}_0, \mathcal{P}_0)$, one has*

$$(\mathcal{L}f)(\xi) \;=\; \int \hat{\xi}(dx)\; c_{0,x}(\xi)\; \nabla_x f(\xi)\;, \qquad \textit{for } \mathcal{P}_0\textit{-a.e. } \xi\;, \tag{32}$$

where $\nabla_x f$ is defined in (17)*, and, moreover,*

$$\langle f, (-\mathcal{L})f\rangle_{\mathcal{P}_0} \;=\; \frac{1}{2}\int \mathcal{P}_0(d\xi)\int \hat{\xi}(dx)\; c_{0,x}(\xi)\;\; (\nabla_x f(\xi))^2\;. \tag{33}$$

Proof. The self-adjointness of $\mathcal{L}$ follows from [RS, Vol.2, Theorem X.1]. Actually, (i) $\mathcal{L}$ is closed as a generator of a strongly continuous semigroup [RS, Vol.2, Chap. X.8]; (ii) $\mathcal{L}$ is symmetric because $\mathcal{T}_t$ is self-adjoint; (iii) the spectrum of $\mathcal{L}$ is included in $(-\infty, 0]$ by contractivity of the semigroup. Note that (iii) also implies that $\mathcal{L}$ is non-positive.

We use the abbreviation L^p for $L^p = L^p(\mathcal{N}_0, \mathcal{P}_0)$, $p = 2$ or ∞. For any $f \in L^\infty$, denote by Λf the function defined by the r.h.s. of (32). Due to Lemma 2, $\mathbf{E}_{\mathcal{P}_0}(\lambda_0^2) < \infty$ and in particular

$$\int \mathcal{P}_0(d\xi)\big|(\Lambda f)(\xi)\big|^2 \;\le\; 4\,\|f\|_\infty^2\,\mathbf{E}_{\mathcal{P}_0}\big(\lambda_0^2\big) < \infty\;,$$

thus implying that $\Lambda : L^\infty \to L^2$ is a well-defined operator. We claim that

$$L^2 - \lim_{t\downarrow 0}\frac{\mathcal{T}_t f - f}{t} \;=\; \Lambda f\;, \qquad \forall\; f \in L^\infty\;. \tag{34}$$

Note that (34) implies that $L^\infty \subset D(\mathcal{L})$ and $\mathcal{L}f = \Lambda f$ for all $f \in L^\infty$. Since moreover $\mathcal{T}_t$ is a contraction and $\mathcal{T}_t\, L^\infty \subset L^\infty$, it then follows from [RS, Vol.2, Theorem X.49] that L_∞ is a core for $\mathcal{L}$ and $\mathcal{L}$ is the closure of Λ. Finally, using (30) in the limit $t \to 0$, by straightforward computations (33) can be derived from (32).

Let us now prove (34). We assume $\xi \in \mathcal{W}$ and we set, for $\xi' \neq \xi$,

$$p_{t,1}(\xi'|\xi) \;:=\; \mathbf{P}_\xi(\xi_t = \xi',\, n_*(t) = 1) \;=\; p_t(\xi'|\xi) - \mathbf{P}_\xi(\xi_t = \xi',\, n_*(t) \ge 2)\;.$$

Thanks to the construction of the dynamics described in Sect. 3.2 and due to the estimate $1 - e^{-u} \le u$, $u \ge 0$, one has for any $x \in \hat{\xi}$ and $x \neq 0$,

$$p_{t,1}(S_x\xi|\xi) \le \tilde{\mathbf{P}}_0^\xi \otimes \mathbf{Q}(\tilde{X}_1^\xi = x,\; T_{0,0}^\xi \le t) = p^\xi(x|0)(1 - e^{-\lambda_0(\xi)t}) \le c_{0,x}(\xi)\, t\;. \tag{35}$$

Let $f \in L^\infty$. In view also of (31) and $\int \hat{\xi}(dx)\; p_t(S_x\xi|\xi) = 1$,

$$\begin{aligned}\Big|\big(\mathcal{T}_t f - f - t\,\Lambda f\big)(\xi)\Big| &= \left|\int \hat{\xi}(dx)\big(f(S_x\xi) - f(\xi)\big)\big(p_t(S_x\xi|\xi) - c_{0,x}(\xi)\,t\big)\right| \\ &\le 2\,\|f\|_\infty\left(\int_{\{x\neq 0\}}\hat{\xi}(dx)\big(p_t(S_x\xi|\xi) - p_{t,1}(S_x\xi|\xi)\big)\right. \\ &\qquad \left. + \int_{\{x\neq 0\}}\hat{\xi}(dx)\big(-p_{t,1}(S_x\xi|\xi) + c_{0,x}(\xi)\,t\big)\right).\end{aligned}$$

The first integral in the second line can be bounded by $\mathbf{P}_\xi(n_*(t) \ge 2)$. The second integral equals

$$-\,\mathbf{P}_\xi(\,n_*(t) = 1\,) + \lambda_0(\xi)\,t \;=\; -1 + e^{-\lambda_0(\xi)t} + \lambda_0(\xi)\,t + \mathbf{P}_\xi(n_*(t) \ge 2)\;.$$

By collecting the above estimates, we get

$$\frac{1}{t^2}\,\mathbf{E}_{\mathcal{P}_0}\big[\big(\mathcal{T}_t f - f - t\Lambda f\big)^2\big] \le \frac{32\,\|f\|_\infty^2}{t^2}\,\mathbf{E}_{\mathcal{P}_0}\Big(\mathbf{P}_\xi^2(n_*(t)\ge 2)\Big) + \frac{8\|f\|_\infty^2}{t^2}\,\mathbf{E}_{\mathcal{P}_0}\Big(\big(-1+e^{-\lambda_0 t}+\lambda_0 t\big)^2\Big). \quad (36)$$

By using the estimate $(e^{-u}-1+u)^2 \le u^3/2$ for $u\ge 0$ and the finiteness of $\mathbf{E}_{\mathcal{P}_0}(\lambda_0^3)$, it is easy to check that the second term in the r.h.s. tends to zero as $t\to 0$. In order to bound the first term, we observe that

$$\begin{aligned}\mathbf{P}_\xi\big(n_*(t)\ge 2\big) &\le \tilde{\mathbf{P}}_0^\xi \otimes \mathbf{Q}(T_{0,0}^\xi \le t,\ T^\xi_{1,\tilde X_1^\xi} \le t)\\ &= \big(1-e^{-\lambda_0(\xi)t}\big)\int \hat\xi(dx)\, p(S_x\xi|\xi)\big(1-e^{-\lambda_0(S_x\xi)t}\big).\end{aligned}$$

Due to the estimate $1-e^{-u}\le u$, this implies the bound

$$\mathbf{P}_\xi(n_*(t)\ge 2) \;\le\; t^2\,\lambda_0(\xi)\int \hat\xi(dx)\, p(S_x\xi|\xi)\lambda_0(S_x\xi) \;=\; t^2\,\lambda_0(\xi)\,\mathbf{E}_{\tilde{\mathbf{P}}_\xi}\big(\lambda_0(\xi_1)\big) \quad (37)$$

Due also to the estimate $1-e^{-u}\le 1$, it is also true that

$$\mathbf{P}_\xi(n_*(t)\ge 2) \;\le\; t\,\mathbf{E}_{\tilde{\mathbf{P}}_\xi}\big(\lambda_0(\xi_1)\big)\,. \quad (38)$$

By multiplying the last two inequalities, and using the stationarity of $\tilde{\mathbf{P}}$, one obtains

$$\begin{aligned}\frac{1}{t^2}\mathbf{E}_{\mathcal{P}_0}\big(\mathbf{P}_\xi^2(n_*(t)\ge 2)\big) &\le t\,\mathbf{E}_{\mathcal{P}_0}\big(\lambda_0(\xi)\big[\mathbf{E}_{\tilde{\mathbf{P}}_\xi}\big(\lambda_0(\xi_1)\big)\big]^2\big)\\ &\le t\,\mathbf{E}_{\mathcal{P}_0}\big(\lambda_0(\xi)\mathbf{E}_{\tilde{\mathbf{P}}_\xi}\big(\lambda_0^2(\xi_1)\big)\big)\\ &= t\,\mathbf{E}_{\mathcal{P}_0}\big(\lambda_0^3\big),\end{aligned}$$

thus implying that the first term on the r.h.s. of (36) goes to 0 as $t\to 0$. □

3.4. Mean forward velocity and infinitesimal square displacement.

Proposition 4. *Let $\mathcal{P}$ satisfy $\rho_{12}<\infty$ and let φ be the $\mathbb{R}^d$-valued function on $\mathcal{N}_0$ and ψ be the function on $\mathcal{N}_0$ with values in the real symmetric $d\times d$ matrices, respectively defined by*

$$\varphi(\xi) \;=\; \int \hat\xi(dx)\, c_{0,x}(\xi)\, x\,, \qquad (a\cdot\psi(\xi)a) \;=\; \int \hat\xi(dx)\, c_{0,x}(\xi)\,(a\cdot x)^2\,. \quad (39)$$

(i) $\varphi(\xi)$ is in $L^2(\mathcal{N}_0,\mathcal{P}_0)$ and is equal to the mean forward velocity given by the convergent L^2-strong limit (21).

(ii) $(a\cdot\psi(\xi)a)$ is in $L^2(\mathcal{N}_0,\mathcal{P}_0)$ and is equal to the infinitesimal mean square displacement given by the convergent L^2-strong limit (22).

We point out that $\varphi(\xi)$ and $\psi(\xi)$ are well defined for $\mathcal{P}_0$ almost all ξ since $\rho_2<\infty$ (see for example the proof of Lemma 2).

Proof. (i) One has

$$\frac{1}{t^2}\int \mathcal{P}_0(d\xi)\big|\mathbf{E}_{\mathbf{P}_\xi}(X_t)-t\,\varphi(\xi)\big|^2 \le \frac{2}{t^2}\int \mathcal{P}_0(d\xi)\,\big|\mathbf{E}_{\mathbf{P}_\xi}\big(X_t\,\chi\big(n_*(t)=1\big)\big)-t\,\varphi(\xi)\big|^2 + \frac{2}{t^2}\int \mathcal{P}_0(d\xi)\,\big|\mathbf{E}_{\mathbf{P}_\xi}\big(X_t\,\chi\big(n_*(t)\ge 2\big)\big)\big|^2\,. \quad (40)$$

We first show that the first term on the r.h.s. vanishes as $t\to 0$. Using the same notation as in the proof of Proposition 3 and invoking (35),

$$\mathbf{E}_{\mathcal{P}_0}\Big(\Big|\mathbf{E}_{\mathbf{P}_\xi}\big(X_t\,\chi\big(n_*(t)=1\big)\big)-t\,\varphi(\xi)\Big|^2\Big) = \mathbf{E}_{\mathcal{P}_0}\Big(\Big|\int_{\{x\neq 0\}}\hat\xi(dx)\big(p_{t,1}(S_x\xi\,|\xi)-t\,c_{0,x}(\xi)\big)\,x\Big|^2\Big)$$

is bounded according to the Cauchy-Schwarz inequality by

$$\mathbf{E}_{\mathcal{P}_0}\Big(\int_{\{x\neq 0\}}\hat\xi(dx)\big(-p_{t,1}(S_x\xi\,|\xi)+t\,c_{0,x}(\xi)\big) \int_{\{y\neq 0\}}\hat\xi(dy)\big(-p_{t,1}(S_y\xi\,|\xi)+t\,c_{0,y}(\xi)\big)|y|^2\Big)\,. \quad (41)$$

Let us denote by $I_1(\xi)$ and $I_2(\xi)$ the (non negative) integrals over $\hat\xi(dx)$ and $\hat\xi(dy)$ respectively. Using the identities of the proof of Proposition 3, the inequality $0\le -1+e^{-u}+u\le u^2$, $u\ge 0$, and (37), we deduce

$$I_1(\xi)=-1+e^{-t\lambda_0(\xi)}+t\lambda_0(\xi)+\mathbf{P}_\xi(n_*(t)\ge 2)\le t^2\lambda_0(\xi)^2+t^2\lambda_0(\xi)\,\mathbf{E}_{\tilde{\mathbf{P}}_\xi}(\lambda_0(\xi_1)).$$

Moreover, $I_2(\xi)\le t\int\hat\xi(dy)c_{0,y}(\xi)|y|^2$. Hence (41) is bounded by

$$t^3\,\Big(\mathbf{E}_{\mathcal{P}_0}\Big(\lambda_0^2(\xi)\int\hat\xi(dy)\,c_{0,y}(\xi)|y|^2\Big)+\mathbf{E}_{\mathcal{P}_0}\Big(\lambda_0(\xi)\mathbf{E}_{\tilde{\mathbf{P}}_\xi}(\lambda_0(\xi_1))\int\hat\xi(dy)\,c_{0,y}(\xi)|y|^2\Big)\Big).$$

As long as $\rho_4<\infty$, the first expression can be bounded by applying Lemma 1(iv) (see the argument leading to Lemma 2). A short calculation shows that the second expression equals

$$\int\mathcal{P}_0(d\xi)\,\int\hat\xi(dx)\,c_{0,x}(\xi)\,\int\hat\xi(dz)\,c_{x,z}(\xi)\,\int\hat\xi(dy)\,c_{0,y}(\xi)\,|y|^2$$

and is therefore bounded if $\rho_4<\infty$ (again by means of Lemma 1(iv)). Resuming the results obtained so far, one gets

$$\frac{1}{t^2}\int\mathcal{P}_0(d\xi)\,\big|\mathbf{E}_{\mathbf{P}_\xi}\big(X_t\,\chi\big(n_*(t)=1\big)\big)-t\,\varphi(\xi)\big|^2 \;=\; \mathcal{O}(t)\,. \quad (42)$$

We now turn to the second term in (40). By Proposition 1, $\mathbf{E}_{\mathcal{P}_0}(\mathbf{E}_{\mathbf{P}_\xi}(|X_t|^\gamma))<\infty$ as long as $0<\gamma<\kappa-3$ whenever $\rho_\kappa<\infty$ for κ integer. By applying twice the Hölder inequality, if $\gamma>2$,

$$\mathbf{E}_{\mathcal{P}_0}\Big(\big|\mathbf{E}_{\mathbf{P}_\xi}\big(X_t\,\chi\big(n_*(t)\ge 2\big)\big)\big|^2\Big)\;\le\;\Big(\mathbf{E}_{\mathcal{P}_0}\Big(\mathbf{E}_{\mathbf{P}_\xi}\big(|X_t|^\gamma\big)\Big)\Big)^{\frac{2}{\gamma}}\Big(\mathbf{E}_{\mathcal{P}_0}\Big(\mathbf{P}_\xi\big(n_*(t)\ge 2\big)^{\frac{2\gamma-2}{\gamma-2}}\Big)\Big)^{1-\frac{2}{\gamma}}\,.$$

Let us take (38) to the power $\gamma/(\gamma-2)$, multiply the result by (37). This yields

$$\mathbf{E}_{\mathcal{P}_0}\Big(\mathbf{P}_\xi\big(n_*(t)\geq 2\big)^{\frac{2\gamma-2}{\gamma-2}}\Big) \;\leq\; t^{\frac{3\gamma-4}{\gamma-2}}\,\mathbf{E}_{\mathcal{P}_0}\Big(\lambda_0(\xi)\,\mathbf{E}_{\tilde{\mathbf{P}}_\xi}\Big(\lambda_0^{\frac{2\gamma-2}{\gamma-2}}(\xi_1)\Big)\Big) = t^{\frac{3\gamma-4}{\gamma-2}}\,\mathbf{E}_{\mathcal{P}_0}\Big(\lambda_0^{\frac{3\gamma-4}{\gamma-2}}\Big)\,.$$

Hence, by Lemma 2, if $\rho_\kappa < \infty$ is satisfied for integer $\kappa > (4\gamma-6)/(\gamma-2)$ and $\gamma < k-3$, there is a finite constant $C>0$ such that

$$\mathbf{E}_{\mathcal{P}_0}\Big(\big|\mathbf{E}_{\mathbf{P}_\xi}\big(X_t\,\chi\big(n_*(t)\geq 2\big)\big)\big|^2\Big) \leq C\,t^{\frac{3\gamma-4}{\gamma}}\,. \tag{43}$$

One concludes the proof by choosing $\gamma > 4$ and by combining (40), (42) and (43), as long as $\kappa > 7$.

(ii) One follows the same strategy. The first term in the equation corresponding to (40) can be dealt with in exactly the same way. In the argument for the second term, $|X_t|$ is replaced by $|X_t|^2$ so that one needs $2\gamma < \kappa - 3$, hence $\kappa > 11$. □

3.5. Proof of Theorem 2. Since all conditions (a)-(d) of Subsect. 3.1 have been checked in the preceding subsections, as already pointed out, one can invoke [DFGW, Theorem 2.2] to conclude that the limit (3) exists and that the rescaled random walk $\underline{Y}^{\xi,\varepsilon}$ converges weakly in $\mathcal{P}_0$–probability to the Brownian motion $\underline{W}_D$. We can now also derive the variational formula (16) from the general expression (23). Let us first quote some general results concerning self–adjoint operators. Let (Ω,μ) be a probability space and denote by $\langle\,.\,,\,.\,\rangle_\mu$ and by $\|.\|_\mu$ the scalar product and the norm on $\mathcal{H}=L^2(\Omega,\mu)$. Let $\mathcal{L}: D(\mathcal{L})\to\mathcal{H}$ be a nonpositive self–adjoint operator with (dense) domain $D(\mathcal{L})\subset\mathcal{H}$ and assume $\mathcal{C}\subset D(\mathcal{L})$ is a core of $\mathcal{L}$. The space $\mathcal{H}_1$ is the completion of $D(|\mathcal{L}|^{1/2})\cap(\mathrm{Ker}(\mathcal{L}))^\perp$ under the norm $\|f\|_1 := \big\||\mathcal{L}|^{1/2}f\big\|_\mu$ for $f\in D(|\mathcal{L}|^{1/2})$, while the dual $\mathcal{H}_{-1}$ of $\mathcal{H}_1$ under $\langle\,.\,,\,.\,\rangle_\mu$ can be identified with the completion of $D(|\mathcal{L}|^{-1/2}) = \mathrm{Ran}(|\mathcal{L}|^{1/2})$ under the $\|.\|_{-1}$-norm defined as $\|\varphi\|_{-1} := \big\||\mathcal{L}|^{-1/2}\varphi\big\|_\mu$ for $\varphi\in D(|\mathcal{L}|^{-1/2})$. Given $\varphi\in\mathcal{H}\cap\mathcal{H}_{-1}$, the dual norm $\|\varphi\|_{-1}$ admits several useful characterizations:

$$\|\varphi\|_{-1}^2 \;=\; \sup_{f\in\mathcal{H}_1\cap\mathcal{H}}\frac{|\langle\varphi,\,f\rangle_\mu|^2}{\|f\|_1^2} \;=\; \sup_{f\in\mathcal{C}\cap(\mathrm{Ker}(\mathcal{L}))^\perp}\frac{|\langle\varphi,\,f\rangle_\mu|^2}{\|f\|_1^2}\,, \tag{44}$$

where the last identity results from the fact that $\mathcal{C}$ is a core for $\mathcal{L}$. Moreover, the identity

$$\|\varphi\|_{-1}^2 \;=\; \sup_{f\in\mathcal{C}}\Big(2\,\langle\varphi,\,f\rangle_\mu - \langle f,\,(-\mathcal{L})f\rangle_\mu\Big) \tag{45}$$

is obtained by using the nonlinearity in f of the expression in the r.h.s. of (45) and observing that $\varphi\in(\mathrm{Ker}(\mathcal{L}))^\perp$. Finally, it follows from spectral calculus that

$$\|\varphi\|_{-1}^2 \;=\; \int_0^\infty dt\,\langle\varphi,\,e^{t\mathcal{L}}\varphi\rangle_\mu\,. \tag{46}$$

In what follows, we extend the definition of $\|\cdot\|_{-1}$ to the whole space $\mathcal{H}$ by setting $\|\varphi\|_{-1}:=\infty$ whenever $\varphi\in\mathcal{H}$ and $\varphi\notin\mathcal{H}_{-1}$. Thanks to this choice, identities (44), (45) and (46) are true for all $\varphi\in\mathcal{H}$.

Invoking (45) and (46), one obtains

$$\int_0^\infty dt\,\Big\langle\varphi\cdot a\,,\;e^{t\mathcal{L}}\,\varphi\cdot a\Big\rangle_{\mathcal{P}_0} = \sup_{f\in L^\infty(\mathcal{N}_0,\mathcal{P}_0)}\Big(\,2\,\langle\varphi\cdot a\,,\;f\rangle_{\mathcal{P}_0} - \langle f,\,(-\mathcal{L})f\rangle_{\mathcal{P}_0}\Big). \tag{47}$$

Using (33), (39) and Lemma 1(i), a short calculation starting from (23) yields (16). □

4. Bound by Cut-off on the Transition Rates

This section and the next ones are devoted to the proof of Theorem 1(ii). In particular, we assume that $\hat{\mathcal{P}}$, $\mathcal{P}$ and ν satisfy the conditions of Theorem 1(ii) although many partial results are true under much weaker conditions. The variational formula (16) is particularly suited in order to derive bounds on the diffusion matrix D. For example, due to the monotonicity of the jump rates $c_{x,y}(\xi)$ in the inverse temperature β, one deduces that the diffusion matrix is a non-increasing function of β. The aim of this section is to obtain more quantitative bounds.

Given an energy $0 \le E_c \le 1$, we define the map $\Phi_c : \mathcal{N} \to \hat{\mathcal{N}} := \mathcal{N}(\mathbb{R}^d)$ as follows:

$$\big(\Phi_c(\xi)\big)(A) \;:=\; \xi(A \times [-E_c, E_c])\;, \qquad A \in \mathcal{B}(\mathbb{R}^d)\;. \tag{48}$$

Note that $\hat{\mathcal{P}}^c := \mathcal{P} \circ \Phi_c^{-1}$ is the distribution of a point process on $\mathbb{R}^d$ with finite intensity $\rho_c := \mathbf{E}_{\hat{\mathcal{P}}^c}(\hat{\xi}(C_1)) \le \mathbf{E}_{\mathcal{P}}(\hat{\xi}(C_1)) = \rho$, and in general

$$\mathbf{E}_{\hat{\mathcal{P}}^c}(\hat{\xi}(C_1)^\kappa) \;\le\; \rho_\kappa\;, \qquad \forall\, \kappa > 0\,. \tag{49}$$

In what follows, we assume that $\rho_c > 0$. It can readily be checked that $\hat{\mathcal{P}}^c$ is an ergodic SSPP on $\mathbb{R}^d$. We write $\hat{\mathcal{P}}_0^c$ for the Palm distribution associated to $\hat{\mathcal{P}}^c$. Note that the distribution $\hat{\mathcal{P}}^c$ is obtained from $\hat{\mathcal{P}}$ by δ_c–thinning with $\delta_c := \nu([-E_c, E_c])$. Thus, $\rho_c = \delta_c\, \rho$. The relation between the Palm distributions $\mathcal{P}_0$ and $\hat{\mathcal{P}}_0^c$ is described in the following lemma.

Lemma 3. *For any Borel set $A \in \mathcal{B}(\hat{\mathcal{N}}_0)$ one has $\hat{\mathcal{P}}_0^c(A) = \rho\, \rho_c^{-1} \mathcal{P}_0(\,|E_0| \le E_c,\ \Phi_c(\xi) \in A\,)$.*

Proof. The assertion is proven by comparing the two following identities obtained from (11):

$$\begin{aligned}
\hat{\mathcal{P}}_0^c(A) \;&=\; \frac{1}{\rho_c} \int_{\hat{\mathcal{N}}} \hat{\mathcal{P}}^c(d\hat{\xi}) \int_{C_1} \hat{\xi}(dx) \chi_A(S_x \hat{\xi})\;, \\
\mathcal{P}_0(\,|E_0| \le E_c,\ \Phi_c(\xi) \in A\,) &= \frac{1}{\rho} \int_{\mathcal{N}} \mathcal{P}(d\xi) \int_{C_1} \hat{\xi}(dx)\, \chi(|E_x| \le E_c)\, \chi_A(\Phi_c(S_x \xi)) \\
&= \frac{1}{\rho} \int_{\mathcal{N}} \mathcal{P}(d\xi) \int_{C_1} \big(\Phi_c(\xi)\big)(dx)\, \chi_A\big(S_x(\Phi_c(\xi))\big)\;.
\end{aligned}$$

□

Proposition 5. *Fix a distance $r_c > 0$ and an energy $0 \le E_c \le 1$ and let $\hat{\mathcal{P}}_0^c$ be as above. Moreover, define*

$$\varphi_c(\hat{\xi}) \;:=\; \int \hat{\xi}(dx)\; \hat{c}_{0,x}\; x\;, \qquad (a \cdot \psi_c(\hat{\xi}) a) \;:=\; \int \hat{\xi}(dx)\; \hat{c}_{0,x}\; (a \cdot x)^2 \tag{50}$$

as functions on $\hat{\mathcal{N}}_0$, where $\hat{c}_{0,x} := \chi(|x| \le r_c)$. Then the diffusion matrix D for the process $(X_t^\xi)_{t \ge 0}$ in Theorem 2 *admits the following lower bound*

$$D \;\ge\; \frac{\rho_c}{\rho}\, e^{-r_c - 4\,\beta\, E_c}\; D_c(r_c, E_c)\;,$$

where

$$(a \cdot D_c(r_c, E_c)\, a) \;:=\; \mathbf{E}_{\hat{\mathcal{P}}_0^c}\Big((a \cdot \psi_c a) \Big) \;-\; 2 \int_0^\infty dt\; \Big\langle \varphi_c \cdot a\,,\; e^{t\mathcal{L}_c}\, \varphi_c \cdot a \Big\rangle_{\hat{\mathcal{P}}_0^c}\;, \tag{51}$$

and $\mathcal{L}_c$ is the unique self–adjoint operator on $L^2(\hat{\mathcal{N}}_0, \hat{\mathcal{P}}_0^c)$ such that

$$(\mathcal{L}_c f)(\hat{\xi}) \;=\; \int \hat{\xi}(dx)\; \hat{c}_{0,x} \nabla_x f(\hat{\xi})\;, \qquad \forall\; f \in L^\infty(\hat{\mathcal{N}}_0, \hat{\mathcal{P}}_0^c)\,. \tag{52}$$

One can prove by the same arguments used in the proof of Proposition 3 that $\mathcal{L}_c$ is well-defined and self–adjoint. Let $\hat{\mathbf{P}}^c$ and $\hat{\mathbf{P}}^c_{\hat{\xi}}$ be the probability measures on the path space $\hat{\Xi} := D([0,\infty), \hat{\mathcal{N}}_0)$ associated to the Markov process with generator $\mathcal{L}_c$ and initial distribution $\hat{\mathcal{P}}_0^c$ and $\delta_{\hat{\xi}}$, respectively, with $\hat{\xi} \in \hat{\mathcal{N}}_0$. One can prove that these Markov processes are well–defined (in particular, $\hat{\mathbf{P}}^c_{\hat{\xi}}$ is well–defined for $\hat{\mathcal{P}}_0^c$–almost all $\hat{\xi}$) and exhibit a realization as jump processes by means of the same arguments used in Sect. 3.2 (note that, for a suitable positive constant c, $\int \hat{\xi}(dx)\hat{c}_{0,x} \le c\, \lambda_0(\xi)$ for any $\xi \in \mathcal{N}_0$, thus allowing to exclude explosion phenomena from the results of Appendix A). Finally, given $\underline{\hat{\xi}} \in \hat{\Xi}$, $X_t(\underline{\hat{\xi}})$ is defined as in (19).

Proof. Note that

$$c_{0,x}(\xi) \;\ge\; e^{-r_c - 4\,\beta\, E_c}\; \tilde{c}_{0,x}(\xi)\;,$$

where

$$\tilde{c}_{x,y}(\xi) \;:=\; \chi\big(|E_x| \le E_c,\; |E_y| \le E_c,\; |x-y| \le r_c\big) \quad , \quad x, y \in \hat{\xi}\;.$$

Then (16) implies that $(a \cdot Da) \;\ge\; e^{-r_c - 4\,\beta\, E_c}\; g(a)$, where

$$g(a) \;:=\; \inf_{f \in L^\infty(\mathcal{N}_0, \mathcal{P}_0)} \int \mathcal{P}_0(d\xi) \int \hat{\xi}(dx)\; \tilde{c}_{0,x}(\xi)\, \Big(a \cdot x \;+\; \nabla_x f(\xi)\Big)^2 \;\ge 0\;.$$

By the same arguments used in the proof of Proposition 3 one can show that there is a unique self–adjoint operator $\tilde{\mathcal{L}}$ on $L^2(\mathcal{N}_0, \mathcal{P}_0)$ such that

$$(\tilde{\mathcal{L}} f)(\xi) \;:=\; \int \hat{\xi}(dx)\; \tilde{c}_{0,x}(\xi)\; \nabla_x f(\xi)\;, \qquad \forall\; f \in L^\infty(\mathcal{N}_0, \mathcal{P}_0).$$

Moreover, $L^\infty(\mathcal{N}_0, \mathcal{P}_0)$ is a core of $\tilde{\mathcal{L}}$ and

$$\langle f, (-\tilde{\mathcal{L}}) f\rangle_{\mathcal{P}_0} = \frac{1}{2} \int \mathcal{P}_0(d\xi) \int \hat{\xi}(dx) \tilde{c}_{0,x}(\xi)\, (\nabla_x f(\xi))^2\;, \quad \forall f \in L^\infty(\mathcal{N}_0, \mathcal{P}_0). \tag{53}$$

Next let us introduce the functions

$$\tilde{\varphi}(\xi) \;=\; \int \hat{\xi}(dx)\; \tilde{c}_{0,x}(\xi)\; x\;, \qquad (a \cdot \tilde{\psi}(\xi) a) \;=\; \int \hat{\xi}(dx)\; \tilde{c}_{0,x}(\xi)\; (a \cdot x)^2\;.$$

Then we obtain by means of straightforward computations and the identities (45), (46) and (53) that

$$g(a) = \mathbf{E}_{\mathcal{P}_0}\Big((a\cdot\tilde{\psi}a)\Big) - 2\,\sup_{f\in L^\infty(\mathcal{N}_0,\mathcal{P}_0)}\Big(2\,\langle\tilde{\varphi}\cdot a,\, f\rangle_{\mathcal{P}_0} - \langle f, (-\tilde{\mathcal{L}})f\rangle_{\mathcal{P}_0}\Big)$$
$$= \mathbf{E}_{\mathcal{P}_0}\Big((a\cdot\tilde{\psi}a)\Big) \;-\; 2\int_0^\infty dt\,\Big\langle\tilde{\varphi}\cdot a\,,\, e^{t\tilde{\mathcal{L}}}\,\tilde{\varphi}\cdot a\Big\rangle_{\mathcal{P}_0}\,.$$

At this point, in order to get (51), it is enough to show that

$$\mathbf{E}_{\mathcal{P}_0}\Big((a\cdot\tilde{\psi}a)\Big) = \delta_c\,\mathbf{E}_{\hat{\mathcal{P}}_0^c}\Big((a\cdot\psi_c a)\Big)\,,$$

and

$$\Big\langle\tilde{\varphi}\cdot a\,,\, e^{t\tilde{\mathcal{L}}}\,\tilde{\varphi}\cdot a\Big\rangle_{\mathcal{P}_0} = \delta_c\,\Big\langle\varphi_c\cdot a\,,\, e^{t\mathcal{L}_c}\,\varphi_c\cdot a\Big\rangle_{\hat{\mathcal{P}}_0^c}\,.$$

This can be derived from Lemma 3 and the following identities, where Φ_c is defined by (48):

$$\tilde{\psi} = \chi(|E_0|\le E_c)\,\psi_c\circ\Phi_c\,,$$
$$\tilde{\varphi} = \chi(|E_0|\le E_c)\,\varphi_c\circ\Phi_c\,,$$
$$\tilde{\mathcal{L}}(f\circ\Phi_c) = \chi(|E_0|\le E_c)\,(\mathcal{L}_c f)\circ\Phi_c\,.$$

□

5. Periodic Approximants and Resistor Networks

In this section, we compare $D_c(r_c, E_c)$ to the diffusion coefficient of adequately defined periodic approximants, which then in turn can be calculated as the conductance of a random resistor network as in [DFGW]. There have been numerous works on periodic approximants; a recent one containing further references is [Owh].

5.1. Random walk on a periodized medium. Let us choose a given direction in $\mathbb{R}^d$, say, the direction parallel to the axis of the first coordinate. Given a fixed configuration $\hat{\xi}\in\hat{\mathcal{N}}$ and $N > r_c$, we define the following subsets of $\mathbb{R}^d$

$$Q_N^{\hat{\xi}} := \mathrm{supp}(\hat{\xi})\cap\check{C}_{2N},\quad \Gamma_N^{\pm} := \mathbb{Z}^d\cap\{x : x^{(1)} = \pm N,\, |x^{(j)}| < N \text{ for } j = 2,\dots,d\},$$
$$\overline{\mathcal{V}}_N^{\hat{\xi}} := Q_N^{\hat{\xi}}\cup\Gamma_N^{+}\cup\Gamma_N^{-},\quad B_N^{\hat{\xi}\pm} := Q_N^{\hat{\xi}}\cap B_N^{\pm},$$

where $\check{C}_{2N} := (-N, N)^d$, $B_N^- := \{x\in\check{C}_{2N} : x^{(1)}\in(-N, -N + r_c]\}$ and $B_N^+ := \{x\in\check{C}_{2N} : x^{(1)}\in[N - r_c, N)\}$.

Next let us introduce a graph $(\overline{\mathcal{V}}_N^{\hat{\xi}}, \overline{\mathcal{E}}_N^{\hat{\xi}})$ with set of vertices $\overline{\mathcal{V}}_N^{\hat{\xi}}$ and set of edges $\overline{\mathcal{E}}_N^{\hat{\xi}}$. Two vertices $x, y\in Q_N^{\hat{\xi}}$ are connected by a non-oriented edge $\{x, y\}\in\overline{\mathcal{E}}_N^{\hat{\xi}}$ if and only if $|x - y|\le r_c$; moreover, all vertices $x\in B_N^{\hat{\xi}+}$ (respectively $x\in B_N^{\hat{\xi}-}$) are connected to all $y\in\Gamma_N^+$ (respectively $y\in\Gamma_N^-$) by an edge $\{x, y\}\in\overline{\mathcal{E}}_N^{\hat{\xi}}$ and the points of $\Gamma_N^{\pm}$ are not connected between themselves.

We now define another graph $(\mathcal{V}_N^{\hat{\xi}}, \mathcal{E}_N^{\hat{\xi}})$ obtained from $(\overline{\mathcal{V}}_N^{\hat{\xi}}, \overline{\mathcal{E}}_N^{\hat{\xi}})$ by identifying the vertices

$$x_- \;=\; (-N, x^{(2)}, \dots, x^{(d)}) \qquad \text{and} \qquad x_+ \;=\; (N, x^{(2)}, \dots, x^{(d)}) \;.$$

Let us write $\pi : \overline{\mathcal{V}}_N^{\hat{\xi}} \to \mathcal{V}_N^{\hat{\xi}}$ for the identification map on the sets of vertices. Hence $\pi(\Gamma_N^-) = \pi(\Gamma_N^+)$ and π restricted to $Q_N^{\hat{\xi}}$ is the identity map. The set $\mathcal{V}_N^{\hat{\xi}} = \pi(\overline{\mathcal{V}}_N^{\hat{\xi}})$ represents the medium periodized along the first coordinate. A vertex $y \in \pi(\Gamma_N^-)$ is connected to all vertices $x \in B_N^{\hat{\xi}+} \cup B_N^{\hat{\xi}-}$ by an edge of $\mathcal{E}_N^{\hat{\xi}}$.

Now a continuous–time random walk with state space $\mathcal{V}_N^{\hat{\xi}}$ and infinitesimal generator $\mathcal{L}_N^{\hat{\xi}}$ is given by

$$\big(\mathcal{L}_N^{\hat{\xi}} f\big)(x) \;=\; \sum_{y \in \mathcal{V}_N^{\hat{\xi}} \,:\, \{x,y\} \in \mathcal{E}_N^{\hat{\xi}}} c(\{x,y\}) \big(f(y) - f(x) \big) \;, \qquad \forall\, x \in \mathcal{V}_N^{\hat{\xi}} \;,$$

where the bond-dependent transition rates $c(\{x, y\})$ are defined for any $\{x, y\} \in \mathcal{E}_N^{\hat{\xi}}$ by

$$c(\{x,y\}) \;=\; \begin{cases} 1 & \text{if } x, y \in Q_N^{\hat{\xi}} \;, \\ \frac{1}{|\Gamma_N^-|} & \text{if } x \in \pi(\Gamma_N^-) \text{ or } y \in \pi(\Gamma_N^-) \;. \end{cases} \tag{54}$$

Clearly the generator $\mathcal{L}_N^{\hat{\xi}}$ is symmetric w.r.t. the uniform distribution $m_N^{\hat{\xi}}$ on $\mathcal{V}_N^{\hat{\xi}}$ given by

$$m_N^{\hat{\xi}} \;=\; \frac{1}{\big|\mathcal{V}_N^{\hat{\xi}}\big|} \sum_{x \in \mathcal{V}_N^{\hat{\xi}}} \delta_x \;.$$

Hence the Markov process with generator $\mathcal{L}_N^{\hat{\xi}}$ and initial distribution $m_N^{\hat{\xi}}$ is reversible. Note that it is not ergodic, however, if there are more than one cluster (equivalence class of edges). In the latter case, the ergodic measures are the uniform distributions on a given cluster and this is sufficient for the present purposes.

We write $\mathbf{P}_N^{\hat{\xi}}$ (respectively $\mathbf{P}_{N,x}^{\hat{\xi}}$) for the probability on the path space $\Omega_N^{\hat{\xi}} = D\big([0,\infty), \mathcal{V}_N^{\hat{\xi}}\big)$ associated to the random walk with initial distribution $m_N^{\hat{\xi}}$ (respectively δ_x) and generator $\mathcal{L}_N^{\hat{\xi}}$.

Let us introduce an antisymmetric function $d_1(x, y)$ on $\mathcal{V}_N^{\hat{\xi}}$ such that

$$d_1(x,y) \;=\; \begin{cases} y^{(1)} - x^{(1)} & \text{if } x, y \in Q_N^{\hat{\xi}} \;, \\ y^{(1)} + N & \text{if } y \in Q_N^{\hat{\xi}},\ y^{(1)} < 0,\ x \in \pi(\Gamma_N^-) \;, \\ y^{(1)} - N & \text{if } y \in Q_N^{\hat{\xi}},\ y^{(1)} > 0,\ x \in \pi(\Gamma_N^-) \;. \end{cases}$$

Finally, given $t \geq 0$, we define the random variable

$$X_{N,t}^{(1)\hat{\xi}}(\underline{\omega}) \;=\; \sum_{s \in [0,t] \,:\, \omega_s \neq \omega_{s-}} d_1(\omega_{s-}, \omega_s) \;,$$

where $(\omega_s)_{s\geq 0} \in \Omega_N^{\hat{\xi}}$. It is the sum of position increments along the first coordinate axis for all jumps occurring in the time interval $[0, t]$. Clearly, $X_{N,t}^{(1)\hat{\xi}}$ gives rise to a time-covariant and antisymmetric family so that, as in Sect. 3, [DFGW, Theorem 2.2] can be used in order to deduce the following result.

Proposition 6. *Given $N \in \mathbb{N}$, $N > r_c$, and $\hat{\xi} \in \hat{\mathcal{N}}$,*

$$\lim_{t\uparrow\infty} \frac{1}{t}\, \mathbf{E}_{\mathbf{P}_N^{\hat{\xi}}}\big((X_{N,t}^{(1)\hat{\xi}})^2\big) \;=\; D_N^{\hat{\xi}}\;,$$

where the diffusion coefficient $D_N^{\hat{\xi}}$ is finite and given by

$$D_N^{\hat{\xi}} \;=\; m_N^{\hat{\xi}}\big(\psi_N^{\hat{\xi}}\big) \;-\; 2\int_0^\infty dt\; \langle \varphi_N^{\hat{\xi}}\,,\, e^{t\mathcal{L}_N^{\hat{\xi}}}\,\varphi_N^{\hat{\xi}}\rangle_{m_N^{\hat{\xi}}}\;, \tag{55}$$

with $\psi_N^{\hat{\xi}}$, $\varphi_N^{\hat{\xi}}$ (scalar) functions on $\mathcal{V}_N^{\hat{\xi}}$ defined as

$$\psi_N^{\hat{\xi}}(x) = \sum_{y\,:\,\{y,x\}\in\mathcal{E}_N^{\hat{\xi}}} c(\{x,y\})\; d_1(x,y)^2, \quad \varphi_N^{\hat{\xi}}(x) = \sum_{y\,:\,\{y,x\}\in\mathcal{E}_N^{\hat{\xi}}} c(\{x,y\})\; d_1(x,y). \tag{56}$$

5.2. Link to periodized medium. Here we show that the diffusion matrix (51) can be bounded from below in terms of the average of the diffusion coefficient associated to the periodized random media. Our proof follows the arguments of [DFGW, Prop. 4.13], but additional technical problems are related to the randomness of geometry (absence of any lattice structure) and possible (albeit integrable) singularities of the mean forward velocity and infinitesimal mean square displacement.

Proposition 7. *Suppose that for $1 \leq p \leq 8$*

$$\lim_{N\uparrow\infty} \frac{\rho_c\, \ell(C_{2N})}{\hat{\xi}(C_{2N}) + a_{2N}} \;=\; 1 \qquad \textit{in} \quad L^p(\hat{\mathcal{N}}, \hat{\mathcal{P}}^c)\;, \tag{57}$$

where $\rho_c := \mathbf{E}_{\hat{\mathcal{P}}^c}(\hat{\xi}(C_1))$ and $a_{2N} := |\Gamma_N^{\pm}| = (2N-1)^{d-1}$. Then, for any $t > 0$,

$$\lim_{N\uparrow\infty} \mathbf{E}_{\hat{\mathcal{P}}^c}\Big(m_N^{\hat{\xi}}\big(\psi_N^{\hat{\xi}}\big)\Big) = \mathbf{E}_{\hat{\mathcal{P}}_0^c}\big(\psi_c^{(11)}\big)\;, \tag{58}$$

$$\lim_{N\uparrow\infty} \mathbf{E}_{\hat{\mathcal{P}}^c}\left(\langle \varphi_N^{\hat{\xi}}\,,\, e^{t\mathcal{L}_N^{\hat{\xi}}}\,\varphi_N^{\hat{\xi}}\rangle_{m_N^{\hat{\xi}}}\right) = \left\langle \varphi_c^{(1)}\,,\, e^{t\mathcal{L}_c}\,\varphi_c^{(1)}\right\rangle_{\hat{\mathcal{P}}_0^c}\;, \tag{59}$$

where $\psi_c^{(11)}$ and $\varphi_c^{(1)}$ are the first diagonal matrix element of the matrix ψ_c and the first component of the vector φ_c introduced in (50)*, respectively.*

Since $D_c(r_c, E_c)$ is given by (51) and is a multiple of the identity (*cf.* Remark 2), the identities (58) and (59) combined with Fatou's Lemma immediately imply:

Corollary 1. *Under the same hypothesis as above,*

$$D_c(r_c, E_c) \;\geq\; \left(\limsup_{N\uparrow\infty} \mathbf{E}_{\hat{\mathcal{P}}^c}\big(D_N^{\hat{\xi}}\big)\right) \mathbf{1}_d\;, \tag{60}$$

where $\mathbf{1}_d$ is the $d \times d$ identity matrix.

Before giving the proof, let us comment on its assumptions. In Sect. 6 we will show that condition (57) is always satisfied. Due to (49), $\rho_p < \infty$ implies $\mathbf{E}_{\hat{\mathcal{P}}^c}(\hat{\xi}(C_1)^p) < \infty$ for any $p > 0$. As $\hat{\mathcal{P}}^c$ is ergodic, this implies the following ergodic theorem, an extension of [DV, Theorem 10.2]. We recall that a convex averaging sequence of sets $\{A_n\}$ in $\mathbb{R}^d$ is a sequence of convex sets such that $A_n \subset A_{n+1}$ and A_n contains a ball of radius r_n with $r_n \to \infty$ as $n \to \infty$.

Lemma 4. *Suppose that $\rho_p < \infty$, $p \geq 1$. Then, given a convex averaging sequence of Borel sets $\{A_n\}$ in $\mathbb{R}^d$,*

$$\frac{\hat{\xi}(A_n)}{\rho_c\,\ell(A_n)} \;\to\; 1 \quad \textit{in} \quad L^p(\hat{\mathcal{N}}, \hat{\mathcal{P}}^c)\,, \qquad \textit{and} \qquad \frac{\hat{\xi}(A_n)}{\rho_c\,\ell(A_n)} \;\to\; 1 \qquad \hat{\mathcal{P}}^c\text{-a.s.}$$

We will also need a bound on $\mathbf{E}_{\hat{\mathcal{P}}^c}((\hat{\xi}(A_n)/\ell(A_n))^p)$, uniformly in n, for a sequence of sets that does not satisfy the assumptions of Lemma 4. To this aim we note that, given a Borel set $B \subset \mathbb{R}^d$ which is a union of k non-overlapping cubes of side 1, one has

$$\mathbf{E}_{\hat{\mathcal{P}}^c}\big(\big(\hat{\xi}(B)/k\big)^p\big) \;\leq\; \mathbf{E}_{\hat{\mathcal{P}}^c}\big(\hat{\xi}(C_1)^p\big) \;\leq\; \rho_p\,, \qquad \forall\; p \geq 1\,. \tag{61}$$

This follows from the stationarity of $\hat{\mathcal{P}}^c$ and the convexity of the function $f(x) = x^p$, $x \geq 0$.

Proof of Proposition 7. Without loss of generality, we assume $r_c = 1$. Note that, since $\hat{\mathcal{P}}$ is stationary with finite intensity ρ_1, one has $\mathcal{P}$-a.s. $\hat{\xi}\,(\partial C_k) = 0$ for all $k \in \mathbb{N}$. In what follows we hence may assume ξ to be as such, thus allowing to simplify notation since $C_{2N} \cap \mathrm{supp}(\hat{\xi}) = \check{C}_{2N} \cap \mathrm{supp}(\hat{\xi})$. A key observation in order to prove (58) and (59) is the following identity, valid for any nonnegative measurable function h defined on $\hat{\mathcal{N}}_0$. It follows easily from (12):

$$\mathbf{E}_{\hat{\mathcal{P}}^c}\left(\int_B \hat{\xi}(dx) h(S_x\hat{\xi})\right) \;=\; \rho_c\,\ell(B)\,\mathbf{E}_{\hat{\mathcal{P}}_0^c}(h), \qquad \forall\;\; B \in \mathcal{B}(\mathbb{R}^d). \tag{62}$$

From this identity we can deduce for any $h \in L^2(\mathcal{N}_0, \hat{\mathcal{P}}_0^c)$ that

$$\lim_{N\uparrow\infty} \mathbf{E}_{\hat{\mathcal{P}}^c}\left(\frac{1}{\hat{\xi}(C_{2N}) + a_{2N}} \int_{C_{2N-2}} \hat{\xi}(dx) h(S_x\hat{\xi})\right) \;=\; \mathbf{E}_{\hat{\mathcal{P}}_0^c}(h)\,. \tag{63}$$

In fact, due to (62), it is enough to show that

$$\mathbf{E}_{\hat{\mathcal{P}}^c}\left(\left(\frac{1}{\hat{\xi}(C_{2N}) + a_{2N}} - \frac{1}{\rho_c \ell(C_{2N-2})}\right) \int_{C_{2N-2}} \hat{\xi}(dx) h(S_x\hat{\xi})\right) \downarrow\; 0\,, \qquad \text{as } N \uparrow \infty. \tag{64}$$

By applying twice the Cauchy-Schwarz inequality and by invoking (62), we obtain

$$\begin{aligned}
&\Big(\text{l.h.s. of (64)}\Big)^2 \\
&\leq \mathbf{E}_{\hat{\mathcal{P}}^c}\left(\left(\frac{\rho_c\,\ell(C_{2N-2})}{\hat{\xi}(C_{2N}) + a_{2N}} - 1\right)^2 \frac{\hat{\xi}(C_{2N-2})}{\rho_c^2\,\ell(C_{2N-2})^2}\right) \\
&\qquad \mathbf{E}_{\hat{\mathcal{P}}^c}\left(\frac{1}{\hat{\xi}(C_{2N-2})}\Big(\int_{C_{2N-2}} \hat{\xi}(dx) h(S_x\hat{\xi})\Big)^2\right) \\
&\leq \mathbf{E}_{\hat{\mathcal{P}}^c}\left(\left(\frac{\rho_c\,\ell(C_{2N-2})}{\hat{\xi}(C_{2N}) + a_{2N}} - 1\right)^2 \frac{\hat{\xi}(C_{2N-2})}{\rho_c\,\ell(C_{2N-2})}\right) \mathbf{E}_{\hat{\mathcal{P}}_0^c}(h^2)\,.
\end{aligned}$$

At this point, (64) follows by applying the Cauchy-Schwarz inequality to the first expectation above and then applying (61) and the limit (57) for $p = 4$.

Let now $h_N^{\hat\xi}$ be a function on $\mathcal{V}_N^{\hat\xi}$ such that for some constant $c > 0$ independent of N,

$$|h_N^{\hat\xi}(x)| \;\le\; c \begin{cases} \hat\xi(\overline{B_1(x)}) & \text{if } x \in Q_N^{\hat\xi}, \\ \frac{|B_N^{\hat\xi}|}{a_{2N}} & \text{otherwise}, \end{cases}$$

where $B_N^{\hat\xi} = B_N^{\hat\xi -} \cup B_N^{\hat\xi +}$ and $\overline{B_1(x)}$ is the closed unit ball centered in x. Note that $\psi_N^{\hat\xi}$ and $\varphi_N^{\hat\xi}$ satisfy this inequality. We claim that the mean boundary contribution vanishes in the limit:

$$\lim_{N\uparrow\infty} \mathbf{E}_{\hat{\mathcal{P}}^c}\Bigg(\frac{1}{\hat\xi(C_{2N}) + a_{2N}} \sum_{x\in\mathcal{V}_N^{\hat\xi}\setminus Q_{N-1}^{\hat\xi}} |h_N^{\hat\xi}(x)|^p\Bigg) \;=\; 0\,, \qquad \text{for } 1 \le p \le 4. \tag{65}$$

In fact, the sum in (65) can be bounded by

$$c^p\, a_{2N} \frac{|B_N^{\hat\xi}|^p}{a_{2N}^p} \;+\; c^p \sum_{x\in Q_N^{\hat\xi}\setminus Q_{N-1}^{\hat\xi}} \Big(\hat\xi(\overline{B_1(x)})\Big)^p\,. \tag{66}$$

By the Cauchy-Schwarz inequality

$$\mathbf{E}_{\hat{\mathcal{P}}^c}\Bigg(\frac{a_{2N}}{\hat\xi(C_{2N}) + a_{2N}}\, \frac{|B_N^{\hat\xi}|^p}{a_{2N}^p}\Bigg) \;\le\; \mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Bigg(\frac{a_{2N}^2}{(\hat\xi(C_{2N}) + a_{2N})^2}\Bigg)\mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Bigg(\frac{|B_N^{\hat\xi}|^{2p}}{a_{2N}^{2p}}\Bigg)\,.$$

The first factor on the r.h.s. is negligible as $N \uparrow \infty$ because of the limit (57) for $p = 2$, while the second factor is bounded, uniformly in N, because of (61). For the second summand in (66), we use twice the Cauchy-Schwarz inequality and invoke (62) to deduce

$$\begin{aligned}\mathbf{E}_{\hat{\mathcal{P}}^c}\Bigg(\frac{1}{\hat\xi(C_{2N}) + a_{2N}} \sum_{x\in Q_N^{\hat\xi}\setminus Q_{N-1}^{\hat\xi}} \Big(\hat\xi\big(\overline{B_1(x)}\big)\Big)^p\Bigg) &\le \mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Bigg(\frac{\hat\xi(C_{2N}\setminus C_{2N-2})}{(\hat\xi(C_{2N}) + a_{2N})^2}\Bigg) \\ &\quad\times \mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Bigg(\frac{1}{\hat\xi(C_{2N}\setminus C_{2N-2})}\Big(\sum_{x\in Q_N^{\hat\xi}\setminus Q_{N-1}^{\hat\xi}} \Big(\hat\xi\big(\overline{B_1(x)}\big)\Big)^p\Big)^2\Bigg) \\ \le \Bigg(\rho_c\, \ell(C_{2N}\setminus C_{2N-2})\mathbf{E}_{\hat{\mathcal{P}}^c}\Bigg(\frac{\hat\xi(C_{2N}\setminus C_{2N-2})}{(\hat\xi(C_{2N}) + a_{2N})^2}\Bigg)\Bigg)^{\frac12} &\mathbf{E}^{\frac12}_{\hat{\mathcal{P}}_0^c}\Bigg(\Big(\hat\xi\big(\overline{B_1(0)}\big)\Big)^{2p}\Bigg)\,.\end{aligned}$$

The last factor is bounded by hypothesis, the first one converges to 0 as $N \uparrow \infty$ because of Lemma 4 and (57).

In order to prove (58) observe that $\psi_c^{(11)}(S_x\hat\xi) = \psi_N^{\hat\xi}(x)$ if $x \in Q_{N-1}^{\hat\xi}$. Therefore we can write

$$m_N^{\hat\xi}(\psi_N^{\hat\xi}) = \frac{1}{\hat\xi(C_{2N}) + a_{2N}} \int_{C_{2N-2}} \hat\xi(dx)\psi_c^{(11)}(S_x\hat\xi) + \frac{1}{\hat\xi(C_{2N}) + a_{2N}} \sum_{x\in\mathcal{V}_N^{\hat\xi}\setminus C_{2N-2}} \psi_N^{\hat\xi}(x)\,.$$

Now (58) follows easily from (63) and (65) with $h_N^{\hat\xi} := \psi_N^{\hat\xi}$. Note that by the same arguments one can prove

$$\lim_{N\uparrow\infty} \mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{ m_N^{\hat\xi}\big[\ |\varphi_N^{\hat\xi}(x)|^p\ \big]\Big\} \ =\ \mathbf{E}_{\hat{\mathcal{P}}_0^c}(|\varphi_c^{(1)}|^p)\ <\ \infty\,, \qquad 1\le p\le 4\,, \tag{67}$$

which will be useful below.

In order to prove (59), we fix $0<\alpha<1$ and set $M = 2N - 2[N^\alpha]$, where $[N^\alpha]$ denotes the integer part of N^α. Moreover, we define the hitting times

$$\tau_N^{\hat\xi}(\underline{\omega}) = \inf\{s\ge 0\,:\,\omega_s\notin C_{2N-2}\}\ ,\quad \underline{\omega}=(\omega_s)_{s\ge 0}\in\Omega_N^{\hat\xi} = D([0,\infty),\mathcal{V}_N^{\hat\xi}). \tag{68}$$

Recall the definitions of the distribution $\hat{\mathbf{P}}^c_{\hat\xi}$, $\mathbf{P}^{\hat\xi}_{N,x}$ and $\mathbf{P}^{\hat\xi}_N$ given in Sects. 4 and 5.1. Thanks to the identity $(e^{t\mathcal{L}_N^{\hat\xi}}\varphi_N^{\hat\xi})(x) = \mathbf{E}_{\mathbf{P}^{\hat\xi}_{N,x}}\big(\varphi_N^{\hat\xi}(\omega_t)\big)$, we can write

$$\mathbf{E}_{\hat{\mathcal{P}}^c}\Big(\langle\varphi_N^{\hat\xi}\,,\ e^{t\mathcal{L}_N^{\hat\xi}}\,\varphi_N^{\hat\xi}\rangle_{m_N^{\hat\xi}}\Big) = \mathbf{E}_{\hat{\mathcal{P}}^c}\big(A_{1,N}^{\hat\xi} + A_{2,N}^{\hat\xi} + A_{3,N}^{\hat\xi}\big),$$

where

$$\begin{aligned}
A_{1,N}^{\hat\xi} &= m_N^{\hat\xi}\Big(\chi(x\notin C_M)\,\varphi_N^{\hat\xi}(x)\,\mathbf{E}_{\mathbf{P}^{\hat\xi}_{N,x}}\big(\varphi_N^{\hat\xi}(\omega_t)\big)\Big),\\
A_{2,N}^{\hat\xi} &= m_N^{\hat\xi}\Big(\chi(x\in C_M)\,\varphi_N^{\hat\xi}(x)\,\mathbf{E}_{\mathbf{P}^{\hat\xi}_{N,x}}\big(\,\chi(\tau_N^{\hat\xi}\le t)\,\varphi_N^{\hat\xi}(\omega_t)\big)\Big),\\
A_{3,N}^{\hat\xi} &= m_N^{\hat\xi}\Big(\chi(x\in C_M)\,\varphi_N^{\hat\xi}(x)\,\mathbf{E}_{\mathbf{P}^{\hat\xi}_{N,x}}\big(\,\chi(\tau_N^{\hat\xi}> t)\,\varphi_N^{\hat\xi}(\omega_t)\big)\Big).
\end{aligned}$$

Then (59) follows from

$$\begin{aligned}
&\lim_{N\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}^c}\big(A_{1,N}^{\hat\xi}\big) = 0,\\
&\lim_{N\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}^c}\big(A_{2,N}^{\hat\xi}\big) = 0,\\
&\lim_{N\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}^c}\big(A_{3,N}^{\hat\xi}\big) = \langle\varphi_c^{(1)}\,,\ e^{t\mathcal{L}_c}\ \varphi_c^{(1)}\rangle_{\hat{\mathcal{P}}_0^c}.
\end{aligned}\tag{69}$$

Let us first prove the first limit in (69). By several applications of Cauchy-Schwarz inequality and due to the identity $\mathbf{P}_N^{\hat\xi} = \int m_N^{\hat\xi}(dx)\mathbf{P}_{N,x}^{\hat\xi}$, we get

$$\begin{aligned}
\big|\mathbf{E}_{\hat{\mathcal{P}}^c}\big(A_{1,N}^{\hat\xi}\big)\big| &\le \mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat\xi}\big(\mathcal{V}_N^{\hat\xi}\setminus C_M\big)\Big\}\mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat\xi}\big[\ \varphi_N^{\hat\xi}(x)^2\mathbf{E}_{\mathbf{P}^{\hat\xi}_{N,x}}\big(\varphi_N^{\hat\xi}(\omega_t)^2\big)\ \big]\Big\}\\
&\le \mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat\xi}\big(\mathcal{V}_N^{\hat\xi}\setminus C_M\big)\Big\}\ \mathbf{E}^{\frac14}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat\xi}\big[\ \varphi_N^{\hat\xi}(x)^4\ \big]\Big\}\ \mathbf{E}^{\frac14}_{\hat{\mathcal{P}}^c}\Big\{\mathbf{E}_{\mathbf{P}^{\hat\xi}_N}\big(\varphi_N^{\hat\xi}(\omega_t)^4\big)\ \big]\Big\}\\
&= \mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat\xi}\big(\mathcal{V}_N^{\hat\xi}\setminus C_M\big)\Big\}\ \mathbf{E}^{\frac12}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat\xi}\big[\ \varphi_N^{\hat\xi}(x)^4\ \big]\Big\}\ ,
\end{aligned}$$

where the last identity follows from the stationarity of $\mathcal{L}_N^{\hat\xi}$ w.r.t. $m_N^{\hat\xi}$. Due to the dominated convergence theorem, the first expectation on the r.h.s. goes to 0, while the second expectation is bounded due to (67).

In order to prove the second limit in (69), we apply twice the Cauchy-Schwarz inequality in order to obtain the bound $\mathbf{E}_{\hat{\mathcal{P}}^c}\big(A^{\hat{\underline{\xi}}}_{2,N}\big)$ by

$$\mathbf{E}^{\frac{1}{2}}_{\hat{\mathcal{P}}^c}\Big\{m^{\hat{\underline{\xi}}}_N\big[\varphi^{\hat{\underline{\xi}}}_N(x)^2\ \big]\Big\}\mathbf{E}^{\frac{1}{4}}_{\hat{\mathcal{P}}^c}\Big\{\mathbf{E}_{\mathbf{P}^{\hat{\underline{\xi}}}_N}\big[\varphi^{\hat{\underline{\xi}}}_N(\omega_t)^4\big]\Big\}\mathbf{E}^{\frac{1}{4}}_{\hat{\mathcal{P}}^c}\Big\{m^{\hat{\underline{\xi}}}_N\big[\chi(x\in C_M)\mathbf{P}^{\hat{\underline{\xi}}}_{N,x}(\tau^{\hat{\underline{\xi}}}_N\le t)\big]\Big\} \tag{70}$$

Again, because of stationarity and (67), the first two factors on the r.h.s. are bounded while the last one converges to 0 due to Lemma 5 below.

Finally we prove the last limit in (69). To this aim, given $\hat{\underline{\xi}}\in\hat{\Xi}=D([0,\infty),\hat{\mathcal{N}}_0)$ and $x\in C_M$, we set

$$\tau_{N,x}(\hat{\underline{\xi}})=\inf\Big\{s\ge 0\,:\,x+X_s(\hat{\underline{\xi}})\notin C_{2N-2}\Big\}\ , \tag{71}$$

where $X_s(\hat{\underline{\xi}})$ is defined as in (19). Note that for $x\in C_M\cap\mathrm{supp}(\hat{\xi})$,

$$\varphi^{\hat{\underline{\xi}}}_N(x)=\varphi^{(1)}_c(S_x\hat{\xi}),\quad \mathbf{E}_{\mathbf{P}^{\hat{\underline{\xi}}}_{N,x}}\big(\chi(\tau^{\hat{\underline{\xi}}}_N>t)\,\varphi^{\hat{\underline{\xi}}}_N(\omega_t)\big)=\mathbf{E}_{\hat{\mathbf{P}}^c_{S_x\hat{\xi}}}\big(\chi(\tau_{N,x}>t)\,\varphi^{(1)}_c(\hat{\xi}_t)\big).$$

Therefore

$$\mathbf{E}_{\hat{\mathcal{P}}^c}\big(A^{\hat{\underline{\xi}}}_{3,N}\big)\ =\ \mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{m^{\hat{\underline{\xi}}}_N\Big[\chi(x\in C_M)\,\varphi^{(1)}_c(S_x\hat{\xi})\,\mathbf{E}_{\hat{\mathbf{P}}^c_{S_x\hat{\xi}}}\big(\chi(\tau_{N,x}>t)\,\varphi^{(1)}_c(\hat{\xi}_t)\big)\Big]\Big\}\,.$$

On the other hand, by applying the Cauchy-Schwarz inequality as in (70) and due to Lemma 5, we obtain

$$\lim_{N\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{m^{\hat{\underline{\xi}}}_N\Big[\chi(x\in C_M)\,|\varphi^{(1)}_c(S_x\hat{\xi})|\,\mathbf{E}_{\hat{\mathbf{P}}^c_{S_x\hat{\xi}}}\big(\chi(\tau_{N,x}\le t)\,|\varphi^{(1)}_c(\hat{\xi}_t)|\big)\Big]\Big\}\ =\ 0\,.$$

The last two identities imply

$$\lim_{N\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}^c}\big(A^{\hat{\underline{\xi}}}_{3,N}\big)\ =\ \lim_{N\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{m^{\hat{\underline{\xi}}}_N\Big[\chi(x\in C_M)\,\varphi^{(1)}_c(S_x\hat{\xi})\,\mathbf{E}_{\hat{\mathbf{P}}^c_{S_x\hat{\xi}}}\big(\varphi^{(1)}_c(\hat{\xi}_t)\big)\Big]\Big\}\,. \tag{72}$$

Observe now that (63) remains valid if the integral is performed on C_M in place of C_{2N-2} (the arguments used in the proof there work also in this case) and the function $h(\hat{\xi})$ is defined as

$$h(\hat{\xi})\ =\ \varphi^{(1)}_c(\hat{\xi})\,\mathbf{E}_{\hat{\mathbf{P}}^c_{\hat{\xi}}}\big(\varphi^{(1)}_c(\hat{\xi}_t)\big)\ =\ \varphi^{(1)}_c(\hat{\xi})\big(e^{t\mathcal{L}_c}\varphi^{(1)}_c\big)(\hat{\xi})\ .$$

Note that $h\in L^2(\hat{\mathcal{N}}_0,\hat{\mathcal{P}}^c_0)$. Therefore we can conclude that the r.h.s. of (72) is equal to $\langle\varphi^{(1)}_c\ ,\ e^{t\mathcal{L}_c}\,\varphi^{(1)}_c\rangle_{\hat{\mathcal{P}}^c_0}$. □

Lemma 5. *Let $\tau^{\hat{\underline{\xi}}}_N$ and $\tau_{N,x}$ be defined as in* (68) *and* (71)*, and let $M=2N-2[N^\alpha]$. Then*

$$\lim_{N\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{m^{\hat{\underline{\xi}}}_N\Big[\chi(x\in C_M)\,\mathbf{P}^{\hat{\underline{\xi}}}_{N,x}\big(\tau^{\hat{\underline{\xi}}}_N\le t\big)\Big]\Big\}=0, \tag{73}$$

$$\lim_{N\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{m^{\hat{\underline{\xi}}}_N\Big[\chi(x\in C_M)\ \hat{\mathbf{P}}^c_{S_x\hat{\underline{\xi}}}\big(\tau_{N,x}\le t\big)\Big]\Big\}=0. \tag{74}$$

Proof. One can check by a coupling argument that the two expectations in (73) and (74) coincide: for each $N \in \mathbb{N}_+$, $\hat{\xi} \in \hat{\mathcal{N}}$ and $x \in C_M \cap \mathrm{supp}(\hat{\xi})$, one can define a probability measure μ on $\Omega_N^{\hat{\xi}} \times \hat{\Xi}$ such that

$$\mu(A \times \hat{\Xi}) = \mathbf{P}_{N,x}^{\hat{\xi}}(A), \quad \mu(\Omega_N^{\hat{\xi}} \times B) = \hat{\mathbf{P}}^c_{S_x\hat{\xi}}(B), \quad \forall A \in \mathcal{B}(\Omega_N^{\hat{\xi}}), \forall B \in \mathcal{B}(\hat{\Xi}),$$

and such that, μ almost surely, $\tau_N^{\hat{\xi}}(\underline{\omega}) = \tau_{N,x}(\underline{\hat{\xi}})$ and $\omega_s = x + X_s(\underline{\hat{\xi}})$ for any $0 \le s < \tau_N^{\hat{\xi}}$. Such a coupling μ implies $\hat{\mathbf{P}}^c_{S_x\hat{\xi}}\big(\tau_{N,x} \le t\big) = \mathbf{P}_{N,x}^{\hat{\xi}}\big(\tau_N^{\hat{\xi}} \le t\big)$. Thus we need to prove only (73). Moreover, without loss of generality, we assume $r_c = 1$.

To this aim let us cover $C_{2N-2} \setminus C_M$ by disjoint cubes $C_{1,i}$ of side 1, $i \in I$, so that $C_{2N-2} \setminus C_M = \cup_{i\in I} C_{1,i}$ (the boundaries of these cubes are suitably chosen for them to be disjoint). Finally, given a positive integer n, we set

$$I_*^n = \{(l_1, \ldots, l_n) \in I^n \,:\, l_j \neq l_k \ \text{ if } j \neq k\}.$$

For paths $\underline{\omega}$ such that $\tau_N^{\hat{\xi}}(\underline{\omega}) < \infty$, let us define $k = k(\underline{\omega})$ as the number of different cubes $C_{1,i}$, $i \in I$, visited by the particle in the time interval $[0, \tau_N^{\hat{\xi}}(\underline{\omega}))$ and moreover we define by induction $(i_1, \ldots, i_k) \in I_*^k$, $(x_1, \ldots, x_k) \in \big(C_{2N-2} \setminus C_M\big)^k$ with $x_j \in C_{1,i_j}$ $\forall j : 1 \le j \le k$, and $(t_1, \ldots, t_k)$ as follows: Let x_1 be the first point reached in $C_{2N-2} \setminus C_M$ and t_1 be the time spent in x_1 before jumping away. The index i_1 is characterized by the requirement that $x_1 \in C_{1,i_1}$. Suppose now that $i_1, \ldots, i_j$, $x_1, \ldots, x_j$ and $t_1, \ldots, t_j$ have been defined and that $j < k$. Then x_{j+1} is the first point in $C_{2N-2} \setminus \big(C_M \cup C_{1,i_1} \cup \cdots \cup C_{1,i_j}\big)$ visited during the time interval $[0, \tau_N^{\hat{\xi}}(\underline{\omega}))$ and t_{j+1} is the time spent at x_{j+1} during such a first visit. Moreover, i_{j+1} is such that $x_{j+1} \in C_{1,i_{j+1}}$.

Now let $T_i^{\hat{\xi}}$, $i \in I$ and $\hat{\xi} \in \hat{\mathcal{N}}$, be a family of independent exponential random variables (all independent from the above random objects) and such that $T_i^{\hat{\xi}}$ has parameter $\hat{\xi}\big(\widetilde{C}_{1,i}\big)$, where

$$\widetilde{C}_{1,i} = \{y \in \mathbb{R}^d \,:\, \mathrm{dist}(y, C_{1,i}) \le 1\}.$$

Since, given $\hat{\xi}$, k and $(x_1, \ldots, x_k)$, t_j $(1 \le j \le k)$ are independent exponential variables and t_j has parameter not larger than $\hat{\xi}\big(\widetilde{C}_{1,i_j}\big)$ and since $k \ge k_{\min} := [N^\alpha] - 1$, we obtain

$$\begin{aligned}
&\mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat{\xi}}\Big[\chi(x \in C_M)\, \mathbf{P}_{N,x}^{\hat{\xi}}\big(\tau_N^{\hat{\xi}} \le t\big)\Big]\Big\} \\
&= \sum_{n=k_{\min}}^{|I|} \sum_{\underline{l}\in I_*^n} \mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat{\xi}}\Big[\chi(x\in C_M) \sum_{\underline{y}\in\prod_{j=1}^n C_{1,l_j}\cap\mathcal{V}_N^{\hat{\xi}}} \mathbf{P}_{N,x}^{\hat{\xi}}\big(\tau_N^{\hat{\xi}} \le t, k=n, x_l=y_l, 1\le l\le n\big)\Big]\Big\} \\
&\le \sum_{n=k_{\min}}^{|I|} \sum_{\underline{l}\in I_*^n} \mathbf{E}_{\hat{\mathcal{P}}^c}\Big\{m_N^{\hat{\xi}}\Big[\chi(x \in C_M)\, \mathbf{P}_{N,x}^{\hat{\xi}}\big(k = n, i_1 = l_1, \ldots, i_n = l_n\big)\Big] \qquad (75) \\
&\qquad \times \mathrm{Prob}\,\big(T_{l_1}^{\hat{\xi}} + \cdots + T_{l_n}^{\hat{\xi}} \le t\big)\Big\},
\end{aligned}$$

where the last inequality follows from the bound

$$\mathbf{P}^{\hat{\xi}}_{N,x}\big(\,\tau^{\hat{\xi}}_N \le t \,|\, k=n, x_1=y_1,\dots,x_n=y_n\big) \le \text{ Prob}\,\big(T^{\hat{\xi}}_{l_1}+\cdots+T^{\hat{\xi}}_{l_n}\le t\big)\,.$$

In order to estimate the probability in the r.h.s., we use an argument similar to that of the proof of Proposition 1 in Appendix C. Let us define $m := \mathbf{E}_{\hat{\mathcal{P}}^c}\big(\hat{\xi}(\widetilde{C}_1)\big)$, where $\widetilde{C}_1 = \{y \in \mathbb{R}^d : \text{dist}(y, C_1) \le 1\}$. Given $\kappa > 0$ and $\underline{l} \in I^n_*$ as above, we define $\mathcal{A} = \mathcal{A}(\kappa, \underline{l})$ as follows

$$\mathcal{A} \;=\; \Big\{\,\hat{\xi}\in\hat{\mathcal{N}} \,:\, \big|\,\{j \,:\, 1\le j\le n \text{ and } \hat{\xi}\big(\widetilde{C}_{1,l_j}\big) > \kappa\, m\,\}\big| > \frac{n}{2}\,\Big\}.$$

Then, by the Chebyshev inequality and the stationarity of $\hat{\mathcal{P}}^c$,

$$\hat{\mathcal{P}}^c\big(\,\mathcal{A}\,\big) \le \frac{2}{n}\mathbf{E}_{\hat{\mathcal{P}}^c}\Big(\big|\,\{j \,:\, 1\le j\le n \text{ and } \hat{\xi}\big(\widetilde{C}_{1,l_j}\big) > \kappa\, m\,\}\big|\Big) \le 2\,\hat{\mathcal{P}}^c\big(\hat{\xi}(\widetilde{C}_1) > \kappa\, m\,\big) \to 0,$$

as $\kappa\to\infty$. Note that the complement $\mathcal{A}^c$ of $\mathcal{A}$ can be written as

$$\mathcal{A}^c \;=\; \Big\{\,\hat{\xi}\in\hat{\mathcal{N}} \,:\, \big|\,\{j \,:\, 1\le j\le n \text{ and } \hat{\xi}\big(\widetilde{C}_{1,l_j}\big) \le \kappa\, m\,\}\big| \ge \big[\,\frac{n}{2}\,\big]_*\,\Big\}\,,$$

where $[n/2]_*$ is defined as $n/2$ for n even and as $(n+1)/2$ for n odd. If $\hat{\xi}\in\mathcal{A}^c$ then at least $\big[\,\frac{n}{2}\,\big]_*$ of the exponential variables $T^{\hat{\xi}}_{l_1},\dots,T^{\hat{\xi}}_{l_n}$ have parameter not larger than $\kappa\, m$. Then, by a coupling argument (*e.g.* Appendix C), we get for all $\hat{\xi}\in\mathcal{A}^c$,

$$\text{Prob}\big(T^{\hat{\xi}}_{l_1}+\cdots+T^{\hat{\xi}}_{l_n}\le t\big) \le\; e^{-\kappa\, m t}\sum_{r=[n/2]_*}^{\infty}\frac{(\kappa\, m\, t)^r}{r!} \;=:\; \phi(\kappa,\,n)\,.$$

Due to the above estimates and since $n \ge k_{\min} = [N^\alpha]-1$, we get

$$\text{r.h.s. of (75)} \;\le\; 2\,\hat{\mathcal{P}}^c\big(\hat{\xi}(\widetilde{C}_1) > \kappa\, m\big) \;+\; \phi(\kappa, N^\alpha)\,.$$

The lemma follows by taking first the limit $N\uparrow\infty$ and then the limit $\kappa\uparrow\infty$. □

5.3. Random resistor networks. We conclude this section by pointing out that the diffusion coefficient $D^{\hat{\xi}}_N$ of the periodized medium can be expressed in terms of the effective conductance of the graph $(\overline{\mathcal{V}}^{\hat{\xi}}_N, \overline{\mathcal{E}}^{\hat{\xi}}_N)$ when assigning suitable bond conductances. More precisely, consider the electrical network given by the graph $(\overline{\mathcal{V}}^{\hat{\xi}}_N, \overline{\mathcal{E}}^{\hat{\xi}}_N)$, where the bond $\{x, y\}\in\overline{\mathcal{E}}^{\hat{\xi}}_N$ has conductivity $c(\{\pi(x),\pi(y)\})$ with $c(\{\cdot,\cdot\})$ defined in (54). Then, the effective conductance $G^{\hat{\xi}}_N$ of this network is defined as the current flowing from Γ^-_N to Γ^+_N when a unit potential difference between Γ^-_N to Γ^+_N is imposed. It can be calculated from Ohm's law and the Kirchhoff rule as follows. Let the electrical potential $V(x)$ vanish on the left border Γ^-_N, be equal to 1 on the right border Γ^+_N, and satisfy:

$$\sum_{y\,:\,\{y,x\}\in\overline{\mathcal{E}}^{\hat{\xi}}_N} c(\{\pi(x),\pi(y)\})\,\big(V(y)-V(x)\big) \;=\; 0 \;\text{ for any }\quad x\in Q^{\hat{\xi}}_N.$$

Then the effective conductance is given by the current flowing through the surfaces $\{x \in [-N, N]^d : x^{(1)} = \pm N\}$:

$$G_N^{\hat{\xi}} \;=\; \sum_{x \in B_N^{\hat{\xi}-}} V(x) \;=\; \sum_{x \in B_N^{\hat{\xi}+}} \big(1 - V(x)\big)\,. \tag{76}$$

By a well-known analogy it is linked to the diffusion coefficient $D_N^{\hat{\xi}}$ (see *e.g.* [DFGW, Prop. 4.15] for a similar proof):

Proposition 8. *One has*

$$D_N^{\hat{\xi}} \;=\; \frac{8\,N^2}{|\overline{\mathcal{V}}_N^{\hat{\xi}}|}\; G_N^{\hat{\xi}}\;. \tag{77}$$

6. Percolation Estimates

Let us set $\mathcal{F}_r := \mathcal{F}_{\mathbb{R}^d \setminus C_r}$ and recall that $\rho_c = \rho\,\delta_c$ with $\delta_c = \nu([-E_c, E_c])$.

6.1. Point density estimates. Here we show how the ergodic properties of Lemma 4 combined with the hypothesis (H1) or (H2) imply (57).

Proposition 9. *Suppose that $\rho_8 < \infty$ and that the hypothesis* (H1) *or* (H2) *holds. For $1 \le p \le 8$,*

$$\lim_{N \uparrow \infty} \frac{\rho_c\,\ell(C_N)}{\hat{\xi}(C_N) + a_N} \;=\; 1 \qquad \textit{in} \quad L^p(\hat{\mathcal{N}}, \hat{\mathcal{P}}^c)\;, \tag{78}$$

where $a_N = (N-1)^{d-1}$.

We will first prove the following criterion.

Lemma 6. *Property* (78) *holds if one has, for some $0 < \rho' < \rho$,*

$$\lim_{N \uparrow \infty}\; N^p\;\; \hat{\mathcal{P}}\left(\hat{\xi}(C_N) \le \rho'\,N^d\right) \;=\; 0\,. \tag{79}$$

Proof. We first check that (79) implies that, for some $0 < \rho'' < \rho'\delta_c$,

$$\lim_{N \uparrow \infty}\; N^p\;\; \hat{\mathcal{P}}^c\left(\hat{\xi}(C_N) \le \rho''\,N^d\right) \;=\; 0\,. \tag{80}$$

If $\delta_c = 1$, this is clearly true so let us suppose that $0 < \delta_c < 1$. Set $\tilde{\delta}_c = 1 - \delta_c$. If $\mathcal{C}_j^k$ denotes the binomial coefficient, we have

$$\begin{aligned}\hat{\mathcal{P}}^c\left(\hat{\xi}(C_N)\le\rho''N^d\right) &= \sum_{k=0}^{[\rho''N^d]}\hat{\mathcal{P}}(\hat{\xi}(C_N)=k)\\ &\quad+\sum_{k=[\rho''N^d]+1}^{\infty}\hat{\mathcal{P}}(\hat{\xi}(C_N)=k)\sum_{j=k-[\rho''N^d]}^{k}\mathcal{C}_j^k\,\tilde{\delta}_c^j\,\delta_c^{k-j}\\ &\le \sum_{k=0}^{[\rho'N^d]}\hat{\mathcal{P}}(\hat{\xi}(C_N)=k)+\sup_{k>[\rho'N^d]}\sum_{j=k-[\rho''N^d]}^{k}\mathcal{C}_j^k\,\tilde{\delta}_c^j\,\delta_c^{k-j}\\ &\le \hat{\mathcal{P}}\left(\hat{\xi}(C_N)\le\rho'N^d\right)+\exp(-c[\rho'N^d](\delta_c-\rho''/\rho')^2),\end{aligned}$$

where the last inequality, given $\rho''<\delta_c\rho'$, follows from a standard large deviation type estimate for Bernoulli variables with some $c>0$. Multiplying by N^p, (79) thus implies (80).

Now set $A_N=\{\hat{\xi}\,:\,\hat{\xi}(C_N)\le\rho''N^d\}$. Then, for some $c'>0$ independent of N,

$$f_N(\hat{\xi})\;:=\;\left|\frac{\rho_c\,\ell(C_N)}{\hat{\xi}(C_N)+a_N}-1\right|^p\;\le\;c'\,\rho_c^p\,N^p\,\chi_{A_N}(\hat{\xi})\;+\;f_N(\hat{\xi})\,\chi_{A_N^c}(\hat{\xi})\;.$$

Integrating w.r.t. $\hat{\mathcal{P}}^c$, the first term vanishes in the limit $N\uparrow\infty$ because of (80). For the second, let us first note that Lemma 4 implies that $\lim_{N\uparrow 0} f_N\chi_{A_N^c}=0$ holds $\hat{\mathcal{P}}^c$-a.s. Furthermore, $|f_N\chi_{A_N^c}|\le c''<\infty$ uniformly in N so that the dominated convergence theorem assures that $\lim_{N\uparrow 0}\mathbf{E}_{\hat{\mathcal{P}}^c}(f_N\chi_{A_N^c})=0$. □

Proof of Proposition 9. Due to Lemma 6 we only need to show that (79) is satisfied for some $\rho'<\rho$. This is trivially true if (H1) holds. Hence let us consider the case where (H2) holds. This implies

$$\left|\mathbf{E}_{\hat{\mathcal{P}}}(f\,|\,\mathcal{F}_{r_2})-\mathbf{E}_{\hat{\mathcal{P}}}(f)\right|\;\le\;\|f\|_\infty\,r_1^d r_2^{d-1}\,h(r_2-r_1)\;,\qquad\hat{\mathcal{P}}\text{-a.s.}\;,\tag{81}$$

where f is a bounded $\mathcal{F}_{C_{r_1}}$–measurable function.

Let C_1^i denote the unit cube centered at $i\in\mathbb{Z}^d$ and $\check{C}_1^i$ be the interior of C_1^i. Let $I_N\subset\mathbb{Z}^d$ be such that $C_N=\cup_{i\in I_N}C_1^i$ and $\check{C}_1^i\cap\check{C}_1^j=\emptyset$ if $i\ne j$. Hence $|I_N|=N^d$. Given $M>0$, set $\tilde{Y}_i(\hat{\xi})=\min\{\hat{\xi}(\check{C}_1^i),\frac{M}{2}\}$ and $Y_i=\tilde{Y}_i-\mathbf{E}_{\hat{\mathcal{P}}}(\tilde{Y}_i)$. Note that Y_i is centered, $\mathcal{F}_{C_1^i}$–measurable and $\|Y_i\|_\infty\le M$. We choose M large enough so that $\tilde{\rho}:=\mathbf{E}_{\hat{\mathcal{P}}}(\tilde{Y}_i)>\rho'$ which is possible because $\lim_{M\uparrow\infty}\mathbf{E}_{\hat{\mathcal{P}}}(\tilde{Y}_i)=\rho>\rho'$. Now

$$\left\{\hat{\xi}(C_N)\le\rho'N^d\right\}\;\subset\;\left\{\sum_{i\in I_N}\tilde{Y}_i(\hat{\xi})\le\rho'N^d\right\}\;\subset\;\left\{\left|\sum_{i\in I_N}Y_i(\hat{\xi})\right|\ge(\tilde{\rho}-\rho')N^d\right\}\;.$$

Hence it is sufficient to show that, for $a>0$,

$$\lim_{N\uparrow\infty}N^p\,\hat{\mathcal{P}}\left(\left|\sum_{i\in I_N}Y_i\right|\ge aN^d\right)\;=\;0\;.\tag{82}$$

By the Chebyshev inequality, one has for any even $q \in \mathbb{N}$:

$$\hat{\mathcal{P}}\left(\left| \sum_{i \in I_N} Y_i \right| \geq a N^d \right) \;\leq\; \frac{1}{a^q \, N^{dq}} \sum_{i_1, \dots, i_q \in I_N} \mathbf{E}_{\hat{\mathcal{P}}} \left(Y_{i_1} \cdots Y_{i_q} \right) \; . \tag{83}$$

We will now bound the sum in the r.h.s. of (83). Let us define the norm $\|x\| = \max\{|x^{(k)}| \, : \, 1 \leq k \leq d\}$ on $\mathbb{R}^d$ (recall that $x^{(k)}$ is the k^{th} component of x) and introduce the notation $\underline{i} = (i_1, \dots, i_q)$, $\underline{I}_N = (I_N)^q$, and $r_j(\underline{i}) = \min\{\|i_j - i_k\| \, : \, k = 1, \dots, q, k \neq j\}$. If $r_1(\underline{i}) = \cdots = r_N(\underline{i}) = 0$, i.e., if each point appears at least twice in $(i_1, \dots, i_q)$, then use the bound $\mathbf{E}_{\hat{\mathcal{P}}}(Y_{i_1} \cdots Y_{i_q}) \leq M^q$. The number of $\underline{i} \in \underline{I}_N$ satisfying this property is at most $c N^{dq/2}$ (here and below c is a varying constant depending only on d and on q). Suppose now that, say, $r_1(\underline{i}) = r \geq 1$. Then the open cubes $\check{C}_1^{i_2}, \dots, \check{C}_1^{i_q}$ are contained in $A := \mathbb{R}^d - C_{2r-1}^{i_1}$ and thus $Y_{i_2}, \dots, Y_{i_q}$ are $\mathcal{F}_A$-measurable. Using the conditional expectation, (81) and the fact that Y_{i_1} is centered and $\|Y_{i_j}\|_\infty \leq M$,

$$\mathbf{E}_{\hat{\mathcal{P}}}(Y_{i_1} \cdots Y_{i_q}) \leq M^{q-1} \, \mathbf{E}_{\hat{\mathcal{P}}}\big(\big| \mathbf{E}_{\hat{\mathcal{P}}}(Y_{i_1} | \mathcal{F}_A) \big| \big) \;\leq\; M^q \, h(2r-2) \, (2r-1)^{d-1} \; .$$

Note that $\mathbf{E}_{\hat{\mathcal{P}}}(Y_{i_1} \cdots Y_{i_q})$ is invariant under permutations of the indices $i_1, \dots, i_q$. Hence

$$\begin{aligned} \sum_{\underline{i} \in \underline{I}_N} \mathbf{E}_{\hat{\mathcal{P}}}(Y_{i_1} \cdots Y_{i_q}) &\leq c \sum_{r=0}^{N} \sum_{\underline{i} \in \underline{K}_N(r)} \mathbf{E}_{\hat{\mathcal{P}}}(Y_{i_1} \cdots Y_{i_q}) \\ &\leq c \, M^q \, N^{dq/2} + c \, M^q \sum_{r=1}^{N} h(2r-2) \, (2r-1)^{d-1} \, |\underline{K}_N(r)| \; , \end{aligned}$$

where $\underline{K}_N(r) = \{\underline{i} \in \underline{I}_N : r_1(\underline{i}) = r, r_2(\underline{i}) \leq r, \dots, r_q(\underline{i}) \leq r\}$. One has

$$|\underline{K}_N(r)| \;\leq\; c \, N^{dq/2} r^{dq/2-1} \; . \tag{84}$$

In fact, on the set of points $i_1, \dots, i_q$ (treated as distinguishable) let us define a graph structure by connecting two points $i \sim j$ with a bond whenever $\|i - j\| \leq r$. We call $\mathcal{G}(\underline{i})$ the resulting graph. Note that each connected component of $\mathcal{G}(\underline{i})$ has cardinality at least 2 whenever $\underline{i} \in \underline{K}_N(r)$, therefore $\mathcal{G}(\underline{i})$ has at most $q/2$ connected components. We claim that, given $1 \leq l \leq q/2$,

$$\big| \{\underline{i} \in \underline{K}_N(r) \, : \, \mathcal{G}(\underline{i}) \text{ has } l \text{ connected components}\} \big| \;\leq\; c(N/r)^{dl} r^{dq-1} \; . \tag{85}$$

In order to prove (85), suppose that the connected component containing i_1 has cardinality k_1, while the other components have cardinality $k_2, \dots, k_l$ respectively. Each component can be built by first choosing one of its points in I_N (there are N^d possible choices), then its neighboring points w.r.t. $\sim$ (for each such neighboring point there are at most cr^d possible choices) and then iteratively adding neighboring points w.r.t. $\sim$. Therefore, the j^{th} component can be built in at most $cN^d r^{d(k_j-1)}$ ways. If $j = 1$, since i_1 has a neighboring point at distance exactly r, the upper bound can be improved by $cN^d r^{d-1+d(k_1-2)}$. Summing over all possible $k_1, \dots, k_l$ such that $k_1 + \cdots + k_l = q$, one gets (85). Since $r \leq N$, (85) implies

$$|\underline{K}_N(r)| \;\leq\; \sum_{l=1}^{q/2} c(N/r)^{dl} r^{dq-1} \;\leq\; c(N/r)^{dq/2} r^{dq-1} \; ,$$

thus concluding the proof of (84). It implies

$$\sum_{\underline{i}\in\underline{I}_N} \mathbf{E}_{\hat{\mathcal{P}}}(Y_{i_1}\cdots Y_{i_q}) \le M^q N^{dq/2}\left(1 + c'\sum_{r=1}^{\infty} h(2r-2)\, r^{dq/2+d-2}\right) . \qquad (86)$$

Provided that $dq > 2p$ and the sum over r converges, that is, if $dq/2 - d \le 8$, we get the result (79) by combining (82), (83), and (86). Choosing for q the smallest even integer larger than $16/d$, (79) is true for $1 \le p \le 8$ and $dq/2 - d \le 8$ as required. □

6.2. Domination. Due to Proposition 9, we may apply the results of Sect. 5 so that combining with Proposition 3,

$$D \;\ge\; \nu([-E_c, E_c])\, e^{-r_c - 4\beta E_c} \limsup_{N\to\infty} \mathbf{E}_{\hat{\mathcal{P}}^c}\left(\frac{8N^2}{|\overline{\mathcal{V}}_N^{\hat{\xi}}|}\, G_N^{\hat{\xi}}\right) . \qquad (87)$$

In order to bound the conductance $G_N^{\hat{\xi}}$ for $N \gg r_c$ from below, we will discretize the space $\mathbb{R}^d$ using cubes of appropriate size and spacing. Given $r_2 \ge r_1 > 0$, let us then consider the following functions on $\hat{\mathcal{N}}$:

$$\sigma_j(\hat{\xi}) \;:=\; \chi\big(\hat{\xi}(C_{r_1} + r_2\, j) > 0\big) , \qquad j \in \mathbb{Z}^d . \qquad (88)$$

They form a random field $\Sigma = (\sigma_j)_{j\in\mathbb{Z}^d}$ on the probability space $(\hat{\mathcal{N}}, \hat{\mathcal{P}}^c)$. If $\hat{\mathcal{P}}$ is a PPP, the σ_j are independent random variables. For a process with finite range correlations, this independence can also be assured by an adequate choice of r_1 and r_2, but in general the σ_j are correlated. The side length r_1 and spacing r_2 are going to be chosen of order $\mathcal{O}(r_c)$ in such a way that all points of neighboring cubes have an euclidean distance less than r_c and they are thus connected by an edge of the graph $(\overline{\mathcal{V}}_N^{\hat{\xi}}, \overline{\mathcal{E}}_N^{\hat{\xi}})$.

Next note that the σ_j take values in $\{0, 1\}$. We shall consider the associated site percolation problem with bonds between nearest neighbors only [Gri]. For this purpose, we shall compare Σ with a random field $Z^p = (z_j^p)_{j\in\mathbb{Z}^d}$ of independent and identically distributed random variables with $\mathrm{Prob}(z_j^p = 1) = p$ and $\mathrm{Prob}(z_j^p = 0) = 1 - p$. In this independent case, it is well-known that there is a critical probability $p_c(d) \in (0, 1)$ such that, if $p > p_c(d)$, there is almost surely a unique infinite cluster, while for $p < p_c(d)$ there is almost surely none [Gri]. We will need somewhat finer estimates for the supercritical regime. Let $|.|$ denote the Euclidean norm in $\mathbb{R}^d$. A left-right crossing (LR-crossing) with length $k-1$ of C_{2N} of a configuration $(z_j^p)_{j\in\mathbb{Z}^d}$ is a sequence of distinct points $y_1, \ldots, y_k$ in $C_{2N} \cap \mathbb{Z}^d$ such that $|y_i - y_{i+1}| = 1$ for $1 \le i < k$, $z_{y_i}^p = 1$ for $1 \le i \le k$, $y_1^{(1)} = -N$, $y_k^{(1)} = N$, $-N < y_i^{(1)} < N$ for $1 < i < k$, and finally $y_i^{(s)} = y_j^{(s)}$ for any $s \ge 3$ and for $1 \le i < j \le k$. Two crossings are called disjoint if all the involved y_j's are distinct. In the same way, one defines disjoint LR-crossings for $(\sigma_j)_{j\in\mathbb{Z}^d}$. Note that this definition of LR-crossings for $d \ge 3$ uses LR-crossings in 2-dimensional slices only. For the random field Z^p, the techniques of [Gri, Sect. 2.6 and 11.3] transposed to site percolation imply that, if $p > p_c(2)$, there are positive constants $a = a(p)$, $b = b(p)$, and $c = c(p)$ such that for all $N \in \mathbb{N}_+$,

$$\mathrm{Prob}\big(Z^p \text{ has less than } bN^{d-1} \text{ disjoint LR–crossings in } C_{2N}\big) \;\le\; c\, e^{-a\,N} . \qquad (89)$$

In order to transpose this result on Z^p to one for Σ, we will use the concept of stochastic dominance [Gri, Sect. 7.4]. One writes $\Sigma \geq_{\rm st} Z^p$ whenever

$$\mathbf{E}_{\hat{\mathcal{P}}^c}(f(\Sigma)) \;\geq\; \mathbf{E}_{\rm Prob}(f(Z^p))\;, \tag{90}$$

for any bounded, increasing, measurable function $f : \{0,1\}^{\mathbb{Z}^d} \to \mathbb{R}$ (recall that a function is increasing if $f((z_j)_{j\in\mathbb{Z}^d}) \geq f((z'_j)_{j\in\mathbb{Z}^d})$ whenever $z_j \geq z'_j$ for all $j \in \mathbb{Z}^d$). As the event on the l.h.s. of (89) is decreasing, $\Sigma \geq_{\rm st} Z^p$ with $p > p_c(2)$ implies that for all $N \in \mathbb{N}_+$,

$$\hat{\mathcal{P}}^c\big((\sigma_j)_{j\in\mathbb{Z}^d} \text{ has less than } bN^{d-1} \text{ disjoint LR–crossings in } C_{2N}\big) \;\leq\; c\, e^{-a\,N}\;. \tag{91}$$

Moreover, let us call the configurations $\hat{\xi}$ in the set on the l.h.s. *N-bad*, those in the complementary set *N-good*. For every N–good configuration $\hat{\xi}$, let us fix a set of at least bN^{d-1} disjoint LR–crossings in C_{2N} for $\big(\sigma_j(\hat{\xi})\big)_{j\in\mathbb{Z}^d}$ and denote it $\mathcal{C}_N(\hat{\xi})$. Given an LR–crossing γ in C_{2N}, we write $L(\gamma)$ for its length. Note that, since the LR–crossings are self–avoiding, $L(\gamma) = |\mathrm{supp}(\gamma)| - 1$ for all $\gamma \in \mathcal{C}_N(\hat{\xi})$. Moreover, since paths in $\mathcal{C}_N(\hat{\xi})$ are disjoint and have support in $C_{2N} \cap \mathbb{Z}^d$, $\sum_{\gamma\in\mathcal{C}_N(\hat{\xi})} |\mathrm{supp}(\gamma)| \leq (2N+1)^d$. The above estimates imply that $\sum_{\gamma\in\mathcal{C}_N(\hat{\xi})} L(\gamma) \leq (2N+1)^d \leq (4N)^d$. In particular, due to the Jensen inequality, for any N–good configuration $\hat{\xi}$,

$$\sum_{\gamma\in\mathcal{C}_N(\hat{\xi})} \frac{1}{L(\gamma)} \;\geq\; \frac{|\mathcal{C}_N(\hat{\xi})|^2}{\sum_{\gamma\in\mathcal{C}_N(\hat{\xi})} L(\gamma)} \;\geq\; \frac{b^2\, N^{d-2}}{4^d}\;. \tag{92}$$

This will allow us to prove a lower bound on (87). Hence we need the following criterion for domination.

Lemma 7. *$\Sigma \geq_{\rm st} Z^p$ holds with $r_1 = r$, $r_2 = 2r$ if $\hat{\mathcal{P}}$ and $r > 0$ satisfy the following: There exists $\rho' > 0$ such that*

$$r^d\, \nu([-E_c, E_c]) \;\geq\; -\frac{\ln(p/2)}{\rho'}\;, \tag{93}$$

and

$$\hat{\mathcal{P}}\Big(\hat{\xi}(C_r) < \rho' r^d \,\Big|\mathcal{F}_{2\,r}\Big) \;\leq\; 1 \,-\, \frac{3\,p}{2}\;, \qquad \hat{\mathcal{P}}\text{–a.s.} \tag{94}$$

Proof. The proof is based on the following criterion [Gri, Sect. 7.4]: if for any finite subset J of $\mathbb{Z}^d$, $i \in \mathbb{Z}^d \setminus J$ and $z_j \in \{0,1\}$ for $j \in J$ satisfying $\hat{\mathcal{P}}^c(\sigma_j = z_j \;\forall j \in J) > 0$, one has

$$\hat{\mathcal{P}}^c(\sigma_i = 1 \,|\, \sigma_j = z_j \;\forall\; j \in J) \;\geq\; p\,, \tag{95}$$

then $\Sigma \geq_{\rm st} Z^p$. Hence let J, i, z_j be as above and set $\tilde{\delta}_c := 1 - \delta_c$ and $J_0 := \{j \in J \,:\, z_j = 0\}$ as well as $J_1 := \{j \in J \,:\, z_j = 1\}$. Moreover, given $\underline{k} \in \mathbb{N}^{J_0}$ and $\underline{s} \in \mathbb{N}_+^{J_1}$, let

$$W(\underline{k}, \underline{s}) := \big\{\hat{\xi} \in \hat{\mathcal{N}} \,:\, \hat{\xi}(C_r + 2rj) = k_j \forall j \in J_0, \hat{\xi}(C_r + 2rj) = s_j \forall j \in J_1\big\}.$$

Then

$$\hat{\mathcal{P}}^c(\sigma_i = 0 \mid \sigma_j = z_j \forall j \in J) = \frac{\sum_{\underline{k}\in\mathbb{N}^{J_0}} \sum_{\underline{s}\in\mathbb{N}_+^{J_1}} \sum_{n\in\mathbb{N}} \hat{\mathcal{P}}\big(\hat{\xi}(C_r+2ri)=n\,,\ W(\underline{k},\underline{s})\big)\, \tilde{\delta}_c^n \prod_{j\in J_0} \tilde{\delta}_c^{k_j} \prod_{j\in J_1}(1-\tilde{\delta}_c^{s_j})}{\sum_{\underline{k}\in\mathbb{N}^{J_0}} \sum_{\underline{s}\in\mathbb{N}_+^{J_1}} \hat{\mathcal{P}}\big(W(\underline{k},\underline{s})\big) \prod_{j\in J_0} \tilde{\delta}_c^{k_j} \prod_{j\in J_1}(1-\tilde{\delta}_c^{s_j})}.$$

Within this, we can, moreover, replace

$$\hat{\mathcal{P}}\big(\hat{\xi}(C_r+2ri)=n\,,\ W(\underline{k},\underline{s})\big) \;=\; \hat{\mathcal{P}}\big(\hat{\xi}(C_r+2ri)=n \mid W(\underline{k},\underline{s})\big)\, \hat{\mathcal{P}}\big(W(\underline{k},\underline{s})\big)\,.$$

Finally, note that $W(\underline{k},\underline{s}) \in \mathcal{F}_A$, where $A = \mathbb{R}^d \setminus \big(C_{2r}+2ri\big)$. As $\tilde{\delta}_c \le e^{-\delta_c}$, we obtain the following bound

$$\sum_{n\in\mathbb{N}} \hat{\mathcal{P}}\big(\hat{\xi}(C_r+2ri)=n \mid W(\underline{k},\underline{s})\big)\tilde{\delta}_c^n \le \hat{\mathcal{P}}\big(\hat{\xi}(C_r+2ri) < \rho' r^d \mid W(\underline{k},\underline{s})\big) \;+\; e^{-\delta_c \rho' r^d}\,.$$

Due to the stationarity of $\hat{\mathcal{P}}$, (93) and (94) imply (95). □

6.3. Proof of Theorem 1(ii). We fix $p > p_c(2)$ and $\rho' < \rho$. Then, given E_c, we choose r_c such that (93) is satisfied, *i.e.* $r_c = c(E_c^{\alpha+1})^{-1/d}$ for some constant c. As $r_c \uparrow \infty$ in the limit of low temperature, we can next check that the condition (94) also holds. This is trivial for a process with a uniform lower bound (4) on the point density. For a mixing point process satisfying (5), one has

$$\hat{\mathcal{P}}\Big(\hat{\xi}(C_r) < \rho' r^d \,\Big|\mathcal{F}_{2r}\Big) \;\le\; \hat{\mathcal{P}}\Big(\hat{\xi}(C_r) < \rho' r^d\Big) \;+\; r^d\,(2r)^{d-1}\,h(r)\,, \qquad \hat{\mathcal{P}}-\text{a.s.}$$

Due to the hypothesis on h, the second term converges to 0 in the limit $r \uparrow \infty$. If $\rho' < \rho$, the first one can be bounded by the Chebychev inequality:

$$\hat{\mathcal{P}}(\hat{\xi}(C_r) \le \rho'\, r^d) \le \hat{\mathcal{P}}\left(\left|\frac{\hat{\xi}(C_r)}{\ell(C_r)} - \rho\right| > \rho - \rho'\right) \le \frac{1}{\rho-\rho'} \int \hat{\mathcal{P}}(d\xi)\, \left|\frac{\hat{\xi}(C_r)}{\ell(C_r)} - \rho\right|.$$

By Lemma 4, the expression on the r.h.s. can be made arbitrarily small by choosing r sufficiently large, thus implying that (94) is satisfied for r sufficiently large. In conclusion, due to Lemma 7, (91) holds for r large enough, *i.e.* temperature low enough. We fix such a value r satisfying (91) and call it r_p.

Consider the variables $(\sigma_j)_{j\in\mathbb{Z}^d}$ defined for $r_1 = r_p, r_2 = 2r_p$ and choose $r_c = (d+8)^{\frac{1}{2}} r_p$. This assures that, if neighboring sites j and j' in $\mathbb{Z}^d$ have $\sigma_j(\hat{\xi}) = \sigma_{j'}(\hat{\xi}) = 1$, then $C_{r_p} + 2jr_p$ and $C_{r_p} + 2j'r_p$ contain each a point and these points are separated by a distance less than r_c. Two neighboring sites j and j' in $\mathbb{Z}^d$ such that $\sigma_j(\hat{\xi}) = \sigma_{j'}(\hat{\xi}) = 1$ define a bond of the site percolation problem. To such a bond one can associate (at least) two points $x \in \operatorname{supp}\hat{\xi} \cap (C_{r_p} + 2jr_p)$ and $y \in \operatorname{supp}\hat{\xi} \cap (C_{r_p} + 2j'r_p)$ separated by a distance less than r_c. Given N integer, we define

$$\hat{N} := \max\big\{n \in \mathbb{N} :\ C_{r_p} + 2r_p j \subset C_{2[r_p N]},\ \ \forall j \in C_{2n} \cap \mathbb{Z}^d\big\}\,.$$

Note that $\hat{N} = \mathcal{O}(N)$. If $j, j' \in C_{2\hat{N}} \cap \mathbb{Z}^d$, then the above associated points x and y are linked by an edge of the graph $(\overline{\mathcal{V}}^{\hat{\xi}}_{[r_pN]}, \overline{\mathcal{E}}^{\hat{\xi}}_{[r_pN]})$ defined in Sect. 5.1. Each LR-crossing of $C_{2\hat{N}}$ for the site percolation problem gives in a natural way a connected path of edges of the graph $(\overline{\mathcal{V}}^{\hat{\xi}}_{[r_pN]}, \overline{\mathcal{E}}^{\hat{\xi}}_{[r_pN]})$ which connects the boundary faces $\Gamma^{\pm}_N$.

For a $\hat{N}$–good configuration $\hat{\xi}$, we now bound the conductance $G^{\hat{\xi}}_{[r_pN]}$ from below. For this purpose, let us consider the random resistor network with vertices $Q^{\hat{\xi}}_{[r_pN]} \cup \{\hat{\Gamma}^+_N, \hat{\Gamma}^-_N\}$, where unit conductances are put on all edges in $\overline{\mathcal{E}}^{\hat{\xi}}_{[r_pN]}$ with vertices in $Q^{\hat{\xi}}_{[r_pN]}$ as well as between the two added boundary points $\hat{\Gamma}^{\pm}_N$ and all points of $B^{\hat{\xi}\pm}_{[r_pN]}$. This new network is obtained from the one of Sect. 5.1 upon placing superconducting wires between all vertices of $\Gamma^+_{[r_pN]}$ and $\Gamma^-_{[r_pN]}$ so that they can be identified with a single point $\hat{\Gamma}^+_N$ and $\hat{\Gamma}^-_N$. The conductance $g^{\hat{\xi}}_N$ of this new network (defined as the current flowing from $\hat{\Gamma}^-_N$ to $\hat{\Gamma}^+_N$ when a unit potential difference is imposed between these two points) is precisely equal to $G^{\hat{\xi}}_{[r_pN]}$ because all points of $\Gamma^{\pm}_{[r_pN]}$ have the same potential (0 or 1 respectively) and each has links to all points of $B^{\hat{\xi}\pm}_{[r_pN]}$ with equal conductances summing up to 1.

In order to bound $g^{\hat{\xi}}_N$ from below, we now invoke Rayleigh's monotonicity law which states that eliminating links (*i.e.* conductances) from the network always lowers its conductance.

For a given $\hat{N}$-good configuration $\hat{\xi}$, we cut all links but those belonging to the family of disjoint paths associated to $\mathcal{C}_{\hat{N}}(\hat{\xi})$. Each of these paths γ connecting $\hat{\Gamma}^+_N$ and $\hat{\Gamma}^-_N$ is self-avoiding and hence has a conductance bounded below by $1/L(\gamma)$. As all the paths of $\mathcal{C}_{\hat{N}}(\hat{\xi})$ are disjoint and they are connecting $\hat{\Gamma}^+_N$ and $\hat{\Gamma}^-_N$ in parallel, $g^{\hat{\xi}}_N$ is the sum of the conductances of all paths and it follows from (92) that $g^{\hat{\xi}}_N \geq c(b)N^{d-2}$ for some positive constant $c(b)$ depending on b. We therefore deduce that

$$\mathbf{E}_{\hat{\mathcal{P}}^c}\left(\frac{[r_pN]^2}{|\overline{\mathcal{V}}^{\hat{\xi}}_{[r_pN]}|} G^{\hat{\xi}}_{[r_pN]}\right) \;\geq\; c(b)\,\mathbf{E}_{\hat{\mathcal{P}}^c}\left(\frac{[r_pN]^2}{|\overline{\mathcal{V}}^{\hat{\xi}}_{[r_pN]}|}\, N^{d-2}\chi(\hat{\xi} \text{ is } \quad \hat{N}\text{–good }\right) .$$

Due to (91) and Proposition 9 the r.h.s. converges to a positive value.

Combining this with the estimate (87) we obtain

$$D \;\geq\; C\;\nu([-E_c, E_c])\; e^{-r_c - 4\beta E_c} \;\geq\; C'\,E_c^{1+\alpha}\,\exp(-cE_c^{-\frac{\alpha+1}{d}} - 4\beta E_c)\;,$$

where C and C' are positive constants. Optimizing the exponent leads to $E_c = c'\,\beta^{-\frac{d}{\alpha+1+d}}$ which completes the proof. □

A. Proof that the Random Walk is Well-Defined

Proposition 10. *Let $\mathcal{P}$ be ergodic with $\rho_2 < \infty$. Then for $\mathcal{P}_0$–almost all $\xi \in \mathcal{N}_0$ and for all $x \in \hat{\xi}$, there exists a unique probability measure $\mathbf{P}^{\xi}_x$ on $\Omega_\xi = D([0,\infty), supp(\hat{\xi}))$ of*

a continuous–time random walk starting at x whose transition probabilities $p_t^\xi(y|x) := \mathbf{P}_x^\xi(X_{s+t}^\xi = y | X_s^\xi = x)$, $x, y \in \hat{\xi}$, $t \geq 0, s \geq 0$ satisfy the infinitesimal conditions (C1) *and* (C2).

Proof. The uniqueness follows from [Bre, Chap. 15]. In order to prove existence, due to the construction described in Sect. 3.2, we only need to prove (27) for $\mathcal{P}_0$–almost all ξ and for any $x \in \hat{\xi}$. According to [Bre, Prop. 15.43], condition (27) is implied by the following one:

$$\tilde{\mathbf{P}}_x^\xi\Big(\sum_{n=0}^{\infty} \frac{1}{\lambda_{\tilde{X}_n^\xi}(\xi)} = \infty\Big) \;=\; 1\;. \tag{96}$$

Due to the identity

$$\tilde{\mathbf{P}}_0^\xi\Big(\sum_{n=1}^{\infty} \frac{1}{\lambda_{\tilde{X}_n^\xi}(\xi)} = \infty \;\Big|\; \tilde{X}_1^\xi = x\Big) \;=\; \tilde{\mathbf{P}}_x^\xi\Big(\sum_{n=0}^{\infty} \frac{1}{\lambda_{\tilde{X}_n^\xi}(\xi)} = \infty\Big)\,, \qquad \forall\, x \in \hat{\xi}\,,$$

the proof will be completed if we can show (96) for $x = 0$ and $\mathcal{P}_0$–almost all ξ and, in particular, if we can show

$$\tilde{\mathbf{P}}\Big(\sum_{n=0}^{\infty} \frac{1}{\lambda_0(\xi_n)} = \infty\Big) \;=\; \int \mathcal{Q}_0(d\xi)\; \tilde{\mathbf{P}}_0^\xi\Big(\sum_{n=0}^{\infty} \frac{1}{\lambda_{\tilde{X}_n^\xi}(\xi)} = \infty\Big) = 1\;,$$

where the distributions $\tilde{\mathbf{P}}$, $\tilde{\mathbf{P}}_0^\xi$, and $\mathcal{Q}_0$ are defined in Sect. 3.3. Due to Proposition 2, $\tilde{\mathbf{P}}$ is ergodic and therefore, according to ergodic theory (see [Ros, Chap. IV]),

$$\lim_{N\uparrow\infty} \frac{1}{N} \sum_{n=0}^{N} \frac{1}{\lambda_0(\xi_n)} \;=\; \mathbf{E}_{\mathcal{Q}_0}\Big(\frac{1}{\lambda_0}\Big) \;=\; \frac{1}{\mathbf{E}_{\mathcal{P}_0}(\lambda_0)}\;, \qquad \tilde{\mathbf{P}}\text{-almost surely,}$$

thus allowing to conclude the proof. □

Remark 3. Explosions are excluded if $\sup_{x\in\hat{\xi}} \lambda_x(\xi) < \infty$ (in such a case (96) is always true), but this simple criterion is typically not satisfied in our case. For instance, for a PPP

$$\sup_{x\in\hat{\xi}} \lambda_x(\xi) \geq e^{-4\beta} \sup_{x\in\hat{\xi}} \sum_{y\in\hat{\xi},\,|y-x|\leq 1} e^{-|x-y|} \geq e^{-4\beta-1} \sup_{x\in\hat{\xi}} \hat{\xi}(C_1 + x) = \infty, \qquad \mathcal{P}_0\text{-a.s.}$$

B. Proof of Lemma 1

Note that the statements (ii) and (iii) of Lemma 1 are proved in [FKAS, Corollary 1.2.11 and Theorem 1.3.9] in dimension $d = 1$. The proof below is valid for any dimension d.

Proof of Lemma 1. (i) Let $h(\xi, \xi') := k(\xi, \xi') - k(\xi', \xi)$. By the definition (11) of the Palm distribution $\mathcal{P}_0$, $\forall N > 0$, $\forall A \in \mathcal{B}(\mathbb{R}^d)$ and for any non negative measurable function f,

$$\int \mathcal{P}_0(d\xi) \int_A \hat{\xi}(dx)\, f(\xi, S_x\xi) = \frac{1}{\rho N^d} \int \mathcal{P}(d\xi) \int_{C_N} \hat{\xi}(dy) \int_{A+y} \hat{\xi}(dx)\, f(S_y\xi, S_x\xi)\;. \tag{97}$$

The antisymmetry of $h(\xi, \xi')$ and the identity above imply

$$\int \mathcal{P}_0(d\xi) \int_{\mathbb{R}^d} \hat{\xi}(dx)\, h(\xi, S_x\xi) = \frac{1}{\rho N^d} \int \mathcal{P}(d\xi) \int_{C_N} \hat{\xi}(dy) \int_{\mathbb{R}^d \setminus C_N} \hat{\xi}(dx) h(S_y\xi, S_x\xi) \quad (98)$$

Let us split the last integral into two integrals over $\mathbb{R}^d \setminus C_{N+\sqrt{N}}$ and over $C_{N+\sqrt{N}} \setminus C_N$. Using (97) again,

$$\begin{aligned} &\frac{1}{\rho N^d} \left| \int \mathcal{P}(d\xi) \int_{C_N} \hat{\xi}(dy) \int_{\mathbb{R}^d \setminus C_{N+\sqrt{N}}} \hat{\xi}(dx) h(S_y\xi, S_x\xi) \right| \\ &\quad \le \int \mathcal{P}_0(d\xi) \int_{\mathbb{R}^d \setminus C_{\sqrt{N}}} \hat{\xi}(dx) \big(|k(\xi, S_x\xi)| + |k(S_x\xi, \xi)| \big) , \end{aligned}$$

which converges to zero as $N \to \infty$ by the dominated convergence theorem. The same holds for

$$\frac{1}{\rho N^d} \left| \int \mathcal{P}(d\xi) \int_{C_N} \hat{\xi}(dy) \int_{C_{N+\sqrt{N}} \setminus C_N} \hat{\xi}(dx)\, h(S_y\xi, S_x\xi) \right| ,$$

since, due to (97), it can be bounded by

$$\begin{aligned} &\frac{1}{\rho N^d} \int \mathcal{P}(d\xi) \int_{C_{N+\sqrt{N}} \setminus C_N} \hat{\xi}(dx) \int_{\mathbb{R}^d} \hat{\xi}(dy) \big(|k(S_y\xi, S_x\xi)| + |k(S_x\xi, S_y\xi)| \big) \\ &= \frac{(N+\sqrt{N})^d - N^d}{N^d} \int \mathcal{P}_0(d\xi) \int_{\mathbb{R}^d} \hat{\xi}(dy) \big(|k(S_y\xi, \xi)| + |k(\xi, S_y\xi)| \big) . \end{aligned}$$

Letting $N \to \infty$ in (98) leads to the result. (ii) Since $\Gamma \in \mathcal{B}(\mathcal{N})$ is translation invariant, one has $\chi_{\Gamma_0}(S_x\xi) = \chi_\Gamma(\xi)$ for all $\xi \in \mathcal{N}$ and $x \in \hat{\xi}$. The above remark together with (11) gives

$$\mathcal{P}_0(\Gamma_0) \;=\; \frac{1}{\rho} \int \mathcal{P}(d\xi) \int_{C_1} \hat{\xi}(dx) \chi_{\Gamma_0}(S_x\xi) \;=\; \frac{1}{\rho} \int_\Gamma \mathcal{P}(d\xi)\, \hat{\xi}(C_1) .$$

Comparing with (1), this yields $\mathcal{P}_0(\Gamma_0) = 1$ if $\mathcal{P}(\Gamma) = 1$. Reciprocally, always due to (1), if $\mathcal{P}_0(\Gamma_0) = 1$, one gets $\hat{\xi}(C_1) = 0$ for $\mathcal{P}$–almost all $\xi \in \mathcal{N} \setminus \Gamma$, and by translation invariance $\xi = 0$ for $\mathcal{P}$–almost all $\xi \in \mathcal{N} \setminus \Gamma$, thus implying that $\mathcal{P}(\Gamma) = 1$. (iii) Let us suppose that $\mathcal{P}_0(A) = \mathcal{P}_0(B) > 0$ and set $\Gamma := \bigcup_{x \in \mathbb{R}^d} S_x B$. This is a translation-invariant Borel subset of $\mathcal{N}$ (see Lemma 8) and $B \subset \Gamma \cap \mathcal{N}_0 \subset A$. In particular, $\mathcal{P}(\Gamma) \in \{0, 1\}$ by the ergodicity of $\mathcal{P}$. Since $\chi_B(S_y\xi) \le \chi_\Gamma(\xi)$ for all $\xi \in \mathcal{N}$ and $y \in \mathbb{R}^d$, it follows from (11) that

$$\mathcal{P}_0(B) \;=\; \frac{1}{\rho} \int_{\mathcal{N}} \mathcal{P}(d\xi) \int_{C_1} \hat{\xi}(dy)\, \chi_B(S_y\xi) \;\le\; \frac{1}{\rho} \int_\Gamma \mathcal{P}(\xi) \hat{\xi}\,(C_1) .$$

Therefore, $\mathcal{P}(\Gamma) = 0$ would imply that $\mathcal{P}_0(B) = 0$, in contradiction with our assumption. Thus $\mathcal{P}(\Gamma) = 1$. But $\Gamma \cap \mathcal{N}_0 \subset A$, therefore the statement follows from (ii). (iv) The thesis follows by observing that (11) implies

$$\mathbf{E}_{\mathcal{P}_0}\Big(\prod_{j=1}^k \hat{\xi}(A_j) \Big) = \frac{1}{\rho} \int_{\mathcal{N}} \mathcal{P}(d\xi) \int_{C_1} \hat{\xi}(dx) \prod_{j=1}^k \hat{\xi}(A_j + x) \le \frac{1}{\rho} \int_{\mathcal{N}} \mathcal{P}(d\xi) \hat{\xi}(C_1) \prod_{j=1}^k \hat{\xi}(\tilde{A}_j)$$

and by applying the estimate $a_1 \cdots a_{k+1} \leq c(k+1)\,(a_1^{k+1}+\cdots+a_{k+1}^{k+1}), a_1, \ldots, a_{k+1} \geq 0$. □

Lemma 8. *Let $A \in \mathcal{B}(\mathcal{N}_0)$. Then $\bigcup_{x\in\mathbb{R}^d} S_x A \in \mathcal{B}(\mathcal{N})$.*

Proof. Let us introduce the following lexicographic ordering on $\mathbb{R}^d$: $x \prec y$ if and only if either $|x| < |y|$ or $|x| = |y|$ and there is k, $1 \leq k \leq d$, such that $x^{(k)} < y^{(k)}$ and $x^{(l)} = y^{(l)}$ for $l < k$ (here $x^{(k)}$ is the k^{th} component of the vector x). Given $\hat{\xi} \in \hat{\mathcal{N}}$, one can then order the support of $\hat{\xi}$ according to $\prec$:

$$\mathrm{supp}(\hat{\xi}) \;=\; \begin{cases} \{y_1(\hat{\xi}), y_2(\hat{\xi}), \ldots, y_N(\hat{\xi})\} & \text{if } N := \hat{\xi}(\mathbb{R}^d) < \infty\,, \\ \{y_j(\hat{\xi})\}_{j\in\mathbb{N}_+} & \text{otherwise}\,, \end{cases}$$

where $y_j \prec y_k$ whenever $j < k$. For any $n \in \mathbb{N}$, let $x_n : \hat{\mathcal{N}} \to \mathbb{R}^d$ then be defined as

$$x_n(\hat{\xi}) \;=\; \begin{cases} y_n(\hat{\xi}) & \text{if } n \leq \hat{\xi}(\mathbb{R}^d)\,, \\ y_N(\hat{\xi}) & \text{if } n > N := \hat{\xi}(\mathbb{R}^d)\,. \end{cases}$$

Using an adequate family of finite disjoint covers of $\mathbb{R}^d$ and the fact that $\hat{\xi} \in \hat{\mathcal{N}} \mapsto \hat{\xi}(B)$ is a Borel function for every Borel set $B \subset \mathbb{R}^d$, one can verify that x_n is a Borel function for each n. Moreover, $\mathrm{supp}(\hat{\xi}) = \{x_n(\hat{\xi}) \,:\, n \in \mathbb{N}\}$ for all $\hat{\xi} \in \hat{\mathcal{N}}$.

Due to the definition of the Borel sets in $\mathcal{N}$ and $\hat{\mathcal{N}}$, the map $\pi : \mathcal{N} \to \hat{\mathcal{N}}$ given by $\pi(\xi) = \hat{\xi}$ is Borel, and by [MKM, Sect. 6.1] the function $F : \mathbb{R}^d \times \mathcal{N} \to \mathcal{N}$ given by $F(x,\xi) = S_x\xi$ is even continuous. Hence we conclude that

$$H_n : \mathcal{N} \to \mathcal{N}_0\,, \qquad H_n(\xi) \;:=\; F\big(x_n(\hat{\xi}), \xi\,\big) \;=\; S_{x_n(\hat{\xi})}\xi\,,$$

is a Borel function. Its restriction $\hat{H}_n : \mathcal{N}_0 \to \mathcal{N}_0$ is then also a Borel function. Now given a Borel subset A of $\mathcal{N}_0$, we conclude that $\Phi(A) := \bigcup_{n=1}^{\infty} \hat{H}_n^{-1}(A)$ is a Borel subset in $\mathcal{N}_0$. One can check that

$$\Phi(A) \;=\; \{\xi \,:\, \xi \;=\; S_x\eta \;\text{ for some } \eta \in A \text{ and } x \in \hat{\eta}\,\}.$$

Since $\mathcal{N}_0$ is a Borel subset of $\mathcal{N}$, it follows that $\Phi(A)$ is a Borel subset of $\mathcal{N}$ as is $H_1^{-1}\big(\,\Phi(A)\,\big)$ since H_1 is a Borel function. The identity

$$H_1^{-1}\big(\,\Phi(A)\,\big) \;=\; \bigcup_{x\in\mathbb{R}^d} S_x A\,,$$

now completes the proof. □

C. Proof of Proposition 1

Proof of Proposition 1. Due to the construction of the dynamics given in Sect. 3.2,

$$\mathbf{E}_{\mathcal{P}_0}\mathbf{E}_{\mathbf{P}_0^\xi}\Big(|X_t^\xi|^\gamma\Big) \;=\; \mathbf{E}_{\mathcal{P}_0}\,\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi\otimes\mathbf{Q}}\big(\,|\tilde{X}^\xi_{n_*^\xi(t)}|^\gamma\,\big)\,.$$

Let $p, q > 1$ be such that $1/p + 1/q = 1$. Due to the Hölder inequality,

$$\mathbf{E}_{\mathcal{P}_0}\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi\otimes\mathbf{Q}}\big(|\tilde{X}^\xi_{n^\xi_*(t)}|^\gamma\big) \;=\; \sum_{n=1}^{\infty}\mathbf{E}_{\mathcal{P}_0}\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi\otimes\mathbf{Q}}\Big(|\tilde{X}^\xi_n|^\gamma\,\chi\big(n^\xi_*(t)\geq 1\big)\chi\big(n^\xi_*(t)=n\big)\Big)$$
$$\leq\; \sum_{n=1}^{\infty}\Big(\mathbf{E}_{\mathcal{P}_0}\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi\otimes\mathbf{Q}}\Big(|\tilde{X}^\xi_n|^{\gamma\,q}\,\chi\big(n^\xi_*(t)\geq 1\big)\Big)\Big)^{\frac{1}{q}}\Big(\mathbf{E}_{\mathcal{P}_0}\big(\,\tilde{\mathbf{P}}_0^\xi\otimes\mathbf{Q}\,(n^\xi_*(t)=n)\big)\Big)^{\frac{1}{p}}.$$

Clearly, $n^\xi_*(t)\geq 1$ means $T^\xi_{0,\tilde{X}^\xi_0}\leq t$. It then follows from the estimate $1-e^{-u}\leq u$, $u\geq 0$, that

$$\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi\otimes\mathbf{Q}}\big(|\tilde{X}^\xi_n|^{\gamma\,q}\,\chi\big(n^\xi_*(t)\geq 1\big)\big)=\big(1-e^{-\lambda_0(\xi)t}\big)\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi}\big(|\tilde{X}^\xi_n|^{\gamma\,q}\big)\leq\lambda_0(\xi)t\;\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi}\big(|\tilde{X}^\xi_n|^{\gamma\,q}\big). \tag{99}$$

We then obtain

$$\mathbf{E}_{\mathcal{P}_0}\,\mathbf{E}_{\mathbf{P}_0^\xi}\big(|X^\xi_t|^\gamma\big)\leq C\sum_{n=1}^{\infty}\Big(\int\mathcal{Q}_0(d\xi)\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi}\Big(|\tilde{X}^\xi_n|^{\gamma\,q}\Big)\Big)^{1/q}$$
$$\Big(\int\mathcal{P}_0(d\xi)\tilde{\mathbf{P}}_0^\xi\otimes\mathbf{Q}\Big(n^\xi_*(t)=n\Big)\Big)^{1/p}, \tag{100}$$

with $C=[t\,\mathbf{E}_{\mathcal{P}_0}(\lambda_0)]^{1/q}$. We claim that there is a (time-independent) constant $C'>0$ such that

$$\int\mathcal{Q}_0(d\xi)\,\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi}\big(|\tilde{X}^\xi_n|^{\gamma\,q}\big)\;\leq\;C'\,n^{\gamma\,q}\,. \tag{101}$$

To show this, let us note first that, given $\tilde{X}^\xi_0=0$, by another application of the Hölder inequality,

$$\big|\tilde{X}^\xi_n\big|^{\gamma\,q}\;=\;\Big|\sum_{m=0}^{n-1}\big(\tilde{X}^\xi_{m+1}-\tilde{X}^\xi_m\big)\Big|^{\gamma\,q}\;\leq\;n^{\gamma\,q-1}\sum_{m=0}^{n-1}\big|\tilde{X}^\xi_{m+1}-\tilde{X}^\xi_m\big|^{\gamma\,q}\,,$$

where it has been assumed that $\gamma\,q>1$. One can derive from the stationarity of $\tilde{\mathbf{P}}$ and Remark 1 that

$$\int\mathcal{Q}_0(d\xi)\,\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi}\big(\big|\tilde{X}^\xi_{n+1}-\tilde{X}^\xi_n\big|^{\gamma\,q}\big)=\int\mathcal{Q}_0(d\xi)\,\mathbf{E}_{\tilde{\mathbf{P}}_0^\xi}\big(\big|\tilde{X}^\xi_1\big|^{\gamma\,q}\big)\;:=\;C'$$

for any $n\in\mathbb{N}$. One concludes the proof of (101) by checking that C' is finite. Actually, by (26), $\mathbf{E}_{\mathcal{P}_0}(\lambda_0)\,C'$ is equal to

$$\int\mathcal{P}_0(d\xi)\int\hat{\xi}(dx)\,c_{0,x}(\xi)|x|^{\gamma\,q}\;\leq\;c\int\mathcal{P}_0(d\xi)\int\hat{\xi}(dx)\,e^{-\frac{|x|}{2}}\,,$$

for a suitable constant c. The r.h.s. can be bounded by means of Lemma 1(iv) and the same argument leading to Lemma 2.

In view of (100) and (101), the proposition will be proved if we can show that the expectation $\mathbf{E}_{\mathcal{P}_0}(\tilde{\mathbf{P}}_0^\xi\otimes\mathbf{Q}(n^\xi_*(t)=n))$ converges to zero more rapidly than $n^{-(\gamma+1)p}$ as

$n \to \infty$. Let us fix $0 < \alpha < 1$. We will show that, if $l > 0$ is such that $\mathbf{E}_{\mathcal{P}_0}(\lambda_0^{l+1}) < \infty$, then

$$\mathbf{E}_{\mathcal{P}_0}\big(\tilde{\mathbf{P}}_0^{\xi} \otimes \mathbf{Q}\big(n_*^{\xi}(t) = n\big)\big) \;=\; \mathcal{O}(n^{-\alpha l})\,. \tag{102}$$

To this end, let us first make a general observation. Let $\lambda > 0$ and let $T_1, \dots, T_k$ be independent exponential variables on some probability space (Ω, μ), with parameters $\lambda_1, \dots, \lambda_k \le \lambda$. Define the random variables $T_j' := (\lambda_j/\lambda)T_j$, $j = 1, \dots, k$. These are independent identically distributed exponential variables with parameter λ. As $T_j' \le T_j$, this shows that

$$\mu\big(T_1 + \cdots + T_k \le t\big) \;\le\; \mu\big(T_1' + \cdots + T_k' \le t\big) \;=\; e^{-\lambda t} \sum_{j=0}^{\infty} \frac{(\lambda t)^{j+k}}{(j+k)!} \;\le\; \frac{(\lambda t)^k}{k!}\,. \tag{103}$$

In order to proceed, for all $\xi \in \mathcal{N}_0$, let us set $B_n^{\xi} := \big\{x \in \hat{\xi} \,:\, \lambda_x(\xi) \le n^{\alpha}\big\}$ as well as

$$A_n^{\xi} \;:=\; \Big\{(\tilde{X}_k^{\xi})_{k\ge 0} \in \tilde{\Omega}_{\xi} \,:\, \exists\, J \subset I_n\,,\; |J| > \frac{n}{2}\,,\; \tilde{X}_j^{\xi} \in B_n^{\xi} \;\; \forall\, j \in J\Big\},$$

where $I_n := \{0, \dots, n-1\}$ and $|J|$ is the cardinality of J. We write $\tilde{\mathbf{P}}_0^{\xi} \otimes \mathbf{Q}\big(n_*^{\xi}(t) = n\big) = g_n(\xi) + h_n(\xi)$ with

$$g_n(\xi) := \tilde{\mathbf{P}}_0^{\xi} \otimes \mathbf{Q}\Big(\big\{n_*^{\xi}(t) = n\big\} \cap A_n^{\xi}\Big), \quad h_n(\xi) := \tilde{\mathbf{P}}_0^{\xi} \otimes \mathbf{Q}\Big(\big\{n_*^{\xi}(t) = n\big\} \cap (A_n^{\xi})^c\Big).$$

We first estimate g_n. Obviously $\{n_*^{\xi}(t) = n\}$ is contained in $\{\sum_{j\in J} T^{\xi}_{j,\tilde{X}_j^{\xi}} \le t\}$. As a result,

$$\begin{aligned} g_n(\xi) \le & \sum_{J\subset I_n, |J|>n/2} \;\sum_{x_0,\dots,x_{n-1}\in\hat{\xi}} \chi\big(x_j \in B_n^{\xi} \;\forall\, j \in J\big)\, \chi\big(x_i \notin B_n^{\xi} \;\forall\, i \in I_n \setminus J\big) \\ & \qquad \tilde{\mathbf{P}}_0^{\xi}\big(\tilde{X}_0^{\xi} = x_0, \dots, \tilde{X}_{n-1}^{\xi} = x_{n-1}\big)\, \mathbf{Q}\Big(\sum_{j\in J} T^{\xi}_{j,x_j} \le t\Big) \\ \le & \max_{k=[n/2]+1,\dots,n-1} \Big\{\frac{(n^{\alpha}\,t)^k}{k!}\Big\}\,. \end{aligned}$$

Thanks to the Stirling formula $k! \sim k^k e^{-k}\sqrt{2\pi k}$ as $k \to \infty$, the last expression can be bounded by a constant times $(2\,e\,t)^{n/2}\, n^{-n(1-\alpha)/2}$ and is thus exponentially small. We now turn to $\mathbf{E}_{\mathcal{P}_0}(h_n)$, $n \ge 1$. Clearly,

$$\tilde{\mathbf{P}}_0^{\xi}\Big(\big(A_n^{\xi}\big)^c\Big) \le \frac{2}{n}\mathbf{E}_{\tilde{\mathbf{P}}_0^{\xi}}\Big(\chi\big(\tilde{X}_0^{\xi} \notin B_n^{\xi}\big) + \cdots + \chi\big(\tilde{X}_{n-1}^{\xi} \notin B_n^{\xi}\big)\Big) = \frac{2}{n}\sum_{m=0}^{n-1} \mathbf{E}_{\tilde{\mathbf{P}}_{\xi}}\big(\lambda_0(\xi_m) > n^{\alpha}\big).$$

By Proposition 2 and invoking Chebyshev's inequality, one obtains for any $l > 0$,

$$\begin{aligned}\mathbf{E}_{\mathcal{P}_0}(h_n) &\le \int \mathcal{P}_0(d\xi)\tilde{\mathbf{P}}_0^{\xi} \otimes \mathbf{Q}\Big(\big\{n_*^{\xi}(t) \ge 1\big\} \cap \big(A_n^{\xi}\big)^c\Big) \le t \int \mathcal{P}_0(d\xi)\,\lambda_0(\xi)\tilde{\mathbf{P}}_0^{\xi}\Big(\big(A_n^{\xi}\big)^c\Big) \\ &\le \frac{2t}{n}\sum_{m=0}^{n-1}\int \mathcal{P}_0(d\xi)\,\lambda_0(\xi)\,\mathbf{E}_{\tilde{\mathbf{P}}_\xi}\big(\lambda_0(\xi_m) > n^{\alpha}\big) = 2t\,\mathbf{E}_{\mathcal{P}_0}\big(\lambda_0\,\chi(\lambda_0 > n^{\alpha})\big) \\ &\le \frac{2t}{n^{\alpha l}}\,\mathbf{E}_{\mathcal{P}_0}\big(\lambda_0^{l+1}\big),\end{aligned}$$

where the second inequality follows from the same argument leading to (99) and the equality follows from the stationarity of $\tilde{\mathbf{P}}$. This proves (102). We may now choose $p = \alpha^{-1} > 1$ arbitrarily close to 1 so that $\gamma q > 1$ and such that one may take for l the smallest integer strictly greater than $\gamma + 1$. For such a choice the sum (100) converges. We can now invoke Lemma 2 to get the result. □

Acknowledgement. We would like to thank A. Bovier, J. Černý, B. Derrida, P. A. Ferrari, D. Gabrielli, A. Ramirez and R. Siegmund-Schultze for very useful comments. The work was supported by the SFB 288, SFB/TR 12 and the Dutch-German Bilateral Research Group "Mathematics of random spatial models from physics and biology".

References

[AHL] Ambegoakar, V., Halperin, B.I., Langer, J.S.: Hopping Conductivity in Disordered Systems. Phys, Rev **B 4**, 2612–2620 (1971)
[BRSW] Bellissard, J., Rebolledo, R., Spehner, D., von Waldenfels, W.: In preparation
[BHZ] Bellissard, J., Hermann, D., Zarrouati, M.: Hull of Aperiodic Solids and Gap Labelling Theorems. In: Directions in Mathematical Quasicrystals, M.B. Baake, R.V. Moody, eds., CRM Monograph Series, Volume **13**, Providence, RI: Amer. Math.Soc., (2000) 207–259
[Bil] Billingsley, P.: Convergence of Probability Measures. New York: Wiley, 1968
[BS] Bolthausen, E., Sznitman, A.-S.: Ten lectures on random media. DMV Seminar **32** Basel: Birkhäuser, 2002
[Bre] Breiman, L.: Probability. Reading, MA: Addison–Wesley, 1953
[DV] Daley, D.J., Vere–Jones, D.: An Introduction to the Theory of Point Processes. New York: Springer, 1988
[DFGW] De Masi, A., Ferrari, P.A., Goldstein, S., Wick, W.D.: An Invariance Principle for Reversible Markov Processes. Applications to Random Motions in Random Environments. J. Stat. Phys. **55**, 787–855 (1989)
[EF] Efros, A.L., Shklovskii, B.I.: Coulomb gap and low temperature conductivity of disordered systems. J. Phys. C: Solid State Phys. **8**, L49–L51 (1975)
[FM] Faggionato, A., Martinelli, F.: Hydrodynamic limit of a disordered lattice gas. Probab. Theory Related Fields **127**, 535–608 (2003)
[FKAS] Franken, P., König, D., Arndt, U., Schmidt, V.: Queues and Point Processes. Berlin: Akadamie-Verlag, 1981
[Gri] Grimmett, G.: Percolation. Second Edition, Grundlehren **321**, Berlin: Springer, 1999
[KV] Kipnis, C., Varadhan, S.R.S.: Central Limit Theorem for Additive Functionals of Reversible Markov Processes and Applications to Simple Exclusion. Commun. Math. Phys. **104**, 1–19 (1986)
[Kal] Kallenberg, O.: Foundations of Modern Probability. Second Edition, New York: Springer-Verlag, 2001
[KLP] Kirsch, W., Lenoble, O., Pastur, L.: On the Mott formula for the a.c. conductivity and binary correlators in the strong localization regime of disordered systems. J. Phys. A: Math. Gen. **36**, 12157–12180 (2003)
[LB] Ladieu, F., Bouchaud, J.-P.: Conductance statistics in small GaAs:Si wires at low temperatures: I. Theoretical analysis: truncated quantum fluctuations in insulating wires. J. Phys. I France **3**, 2311–2320 (1993)
[Mar] Martinelli, F.: Lectures on Glauber dynamics for discrete spin models. Lecture Notes in Mathematics, Vol. **1717**, Berlin-Heidelberg-Newyork: Springer, 2000

[MKM] Matthes, K., Kerstan, J., Mecke, J.: Infinitely Divisible Point Processes. Wiley Series in Probability and Mathematical Physics, Newyork: Wiley, 1978
[MR] Meester, R., Roy, R.: Continuum Percolation. Cambridge: Cambridge University Press, 1996
[MA] Miller, A., Abrahams, E.: Impurity Conduction at Low Concentrations. Phys. Rev. **120**, 745–755 (1960)
[Min] Minami, N.: Local fluctuation of the spectrum of a multidimensional Anderson tight binding model. Commun. Math. Phys. **177**, 709–725 (1996)
[Mot] Mott, N.F.: J. Non-Crystal. Solids **1**, 1 (1968); N. F. Mott, Phil. Mag **19**, 835 (1969); Mott, N.F., Davis, E.A.: Electronic Processes in Non-Crystaline Materials. New York: Oxford University Press, 1979
[Owh] Owhadi, H.: Approximation of the effective conductivity of ergodic media by periodization. Probab. Theory Related Fields **125**, 225–258, (2003)
[Qua] Quastel, J.: Diffusion in Disordered Media. In: Funaki, T., Woyczinky, W., eds., Proceedings on stochastic method for nonlinear P.D.E., IMA volumes in Mathematics **77**, New York: Springer Verlag, 1995, pp. 65–79
[RS] Reed, M., Simon, B.: Methods of Modern Mathematical Physics I-IV. San Diego: Academic Press, 1980
[Ros] Rosenblatt, M.: Markov Processes. Structure and Asymptotic Behavior. Grundlehren **184**, Berlin: Springer, 1971
[SE] Shklovskii, B., Efros, A.L.: Electronic Properties of Doped Semiconductors. Berlin: Springer, 1984
[Spe] Spehner, D.: Contributions à la théorie du transport électronique dissipatif dans les solides apériodiques. PhD Thesis, Toulouse, 2000
[Spo] Spohn, H.: Large Scale Dynamics of Interacting Particles. Berlin: Springer, 1991
[Tho] Thorisson, H.: Coupling, Stationarity, and Regeneration. New York: Springer, 2000

Communicated by M. Aizenman